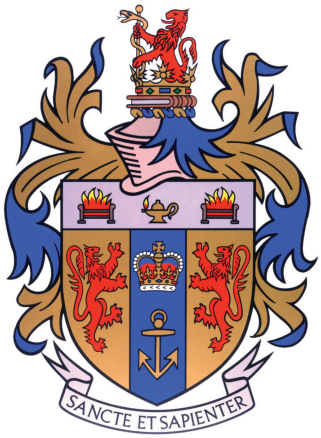
Investigating the Snow Crystal
A Computational Study of the Molecular Mechanisms of Hexagonal Ice Growth

Fulford, Maxwell

Awarding institution:
King's College London

Investigating the Snow Crystal: A Computational Study of the Molecular Mechanisms of Hexagonal Ice Growth



King's College London

Maxwell Fulford

School of Natural & Mathematical Sciences

A thesis presented for the degree of
Doctor of Philosophy

March 2017

Abstract

Hexagonal Ice (I_h) plays a number of essential roles in controlling and maintaining the natural environment on Earth. It is the predominant crystalline form of ice and permanently covers at least 10% of all land. Within the atmosphere ice forms cirrus clouds which are involved in the greenhouse effect. The altitude and thickness of the clouds controls the balance of solar heating and infrared cooling from Earth to space, and is governed by the sizes and shapes of the constituent ice crystals. The shape of vapour grown ice crystals depends strongly on temperature and vapour pressure. At low vapour pressures, simple hexagonal prisms form with an aspect ratio that depends on the relative rate of growth of its two crystallographic surfaces - the basal and prism surface. The competing action of the two rates results in thin hexagonal plates at very low temperatures and elongated prisms at higher temperatures. Curiously, as temperature is increased further, thin plates form once again. The reasons why growth occurs preferentially along the basal and prism surfaces at different temperatures remain unknown. Progress in climate science is hindered by an incomplete grasp of the underlying mechanisms involved in ice growth in cirrus clouds. A comprehensive understanding of ice crystal growth is essential if we are to overcome the colossal climate change challenges we face.

In this project we investigate the basal and secondary prism surfaces of hexagonal ice at a range of temperatures between 240 K and 270 K using molecular dynamics and metadynamics. Metadynamics is an enhanced sampling technique that enables efficient sampling of rare events and yields an estimate of the free energy as a function of selected collective variables. Our simulations show the formation of a disordered quasi-liquid layer (QLL) on the surfaces of ice. The QLL mediates crystal growth and has a thickness which varies with temperature. We investigate how the ice/QLL and QLL/vapour interfaces influence the water adsorption potential, surface diffusion properties and growth shape. Our findings reveal that the outer surface structure at the QLL/vapour interface depends weakly on the underlying crystal lattice. The desorption energy costs are equivalent for the basal and prism surfaces, however, the crystal lattice impacts the dynamics within the QLL. Our results show that there are distinct diffusion energy barriers at the basal and prismatic QLL/ice interfaces. The implication of our work is that the QLL/ice interface is key to the overall kinetics of ice growth in vapour whereas the QLL/vapour interface plays a secondary role.

Acknowledgements

I would firstly like to thank my supervisor, Prof. Carla Molteni, for her academic supervision, sharing her wealth of experience and her personal support throughout my PhD. I am especially grateful for the detailed feedback I received in writing this thesis. It has been a pleasure to spend three and a half years working at King's College London and I would also like to thank all members of staff for their dedication in supporting the students on a daily basis.

I would like to express my gratitude to Engineering and Physical Sciences Research Council for the funding which allowed me to complete this work. I am grateful to the Royal Society for a travel grant which gave me the opportunity to spend time working in the group of Prof. Michele Parrinello in Lugano. I would like to extend my sincere thanks to Dr. Matteo Salvalaglio who I first met in Lugano for his guidance and help throughout my PhD.

I would like to thank my friends and colleagues at King's. To Dan, who started and finished his PhD on practically the same days as me, for his support, help with problem solving and friendship. To Federico, my neighbour, for his great company and help during his time at King's, and to Maria-Celeste who has helped to keep me motivated and sane as I staggered towards the finish line.

I would like to thank my parents, Tim and Sima, and sister Leila for their patience and unwavering support throughout my academic life. Finally, I would like to thank Rosie who has become an expert on the quasi-liquid layer of ice, for her continued support and encouragement throughout my PhD. I am very grateful for the countless hours of proof reading, care and understanding during the hectic final months, and her belief in everything I do.

Contents

List of Figures

Chapter 1

Introduction

Earth's environment depends on two delicately balanced equilibria [1]. Firstly, the ratio between the radiation received from the Sun and the radiation reflected or re-radiated from Earth. And secondly, the evaporation of water from the oceans which leads to the subsequent flow of ice sheets back into the oceans [1]. Ice in the sky, sea and on land plays a vital role in both these equilibria.

Within the atmosphere, ice forms cirrus clouds which permanently cover 30% of the Earth's surface [2]. These clouds are involved in the greenhouse effect and control the balance of solar heating and infrared radiation cooling to space [3]. The sizes and shapes of the constituent ice crystals within the clouds influence the radiative properties [4, 5]. On Earth, 10% of land is permanently covered by ice or snow [6] which insulates the ground and controls the global average temperature [1]. At the poles, ice caps reflect up to 90% of the Sun's radiation [6]. In the seas, on average 7% of all oceanic surfaces are frozen, which influences oceanic currents and the exchange of gasses with seawater [6]. Above the seas, cirrus clouds concentrate airborne chemicals and provide the environment for atmospheric chemical reactions to occur, such as ozone depleting reactions [6]. In this instance, the presence of a disordered "quasi-liquid" layer (QLL) on the surface of ice enables the reactions [7]. However, the underlying molecular mechanisms of ice growth remain for the most part unknown. This lack of knowledge is hindering progress in atmospheric and climate science.

1.1 Ice Growth in Vapour

The life cycle of a crystallising vapour molecule involves transportation to the crystal surface, adsorption onto the surface, the release of latent heat during adsorption, migration across the crystal surface to a growth site, and the release of latent heat at the growth site as the water molecule becomes incorporated into the ice lattice. Each of these steps is reversible and in competition with sublimation. Macroscopically, crystallisation involves the transfer of mass and heat and the combination of molecular and convective diffusion. The latent heat released during crystallisation is proportional to the amount of vapour deposited. Growth at equilibrium dissipates the latent heat at the rate it is released and maintains the surface temperature and surrounding vapour density. The role and influence of these complex processes on ice growth remain uncertain.

The pioneering work in the 1950s by Nakay [8], Kobayashi [9] and Hallett [10] led to the detailed morphology diagram in Figure 1.1 which illustrates the shape ice crystals form when grown at different temperatures and supersaturations or vapour pressures. The experiments reveal a primary dependence of the crystal habit on temperature and a secondary dependence on supersaturation. When the ambient environment of a growing crystal is altered, the experiments show that the crystal growth habit immediately transitions to that of the new environment. The experiments conclude that the crystal shape is a product of the ambient growth environment and not of the nucleation environment of the seed crystal. The morphology diagram reveals two transitions from thin plates to elongated prisms depending on the relative rate of growth of the exposed basal and prism crystallographic faces of hexagonal ice. Below $\sim -8\,^{\circ}\mathrm{C}$, the prism surface grows fastest and thin plates form. Between $\sim -8\,^{\circ}\mathrm{C}$ and $\sim -3\,^{\circ}\mathrm{C}$, the basal surface grows fastest and elongated prisms form. Above $\sim -3\,^{\circ}\mathrm{C}$, the prism surface grows fastest once again and thin plates form. However, the underlying processes which influence and control the growth shape are far from understood and the dominant factors are unclear.

A comprehensive understanding of ice crystal growth is required if we are to overcome the significant climate change challenges that the world faces. This must

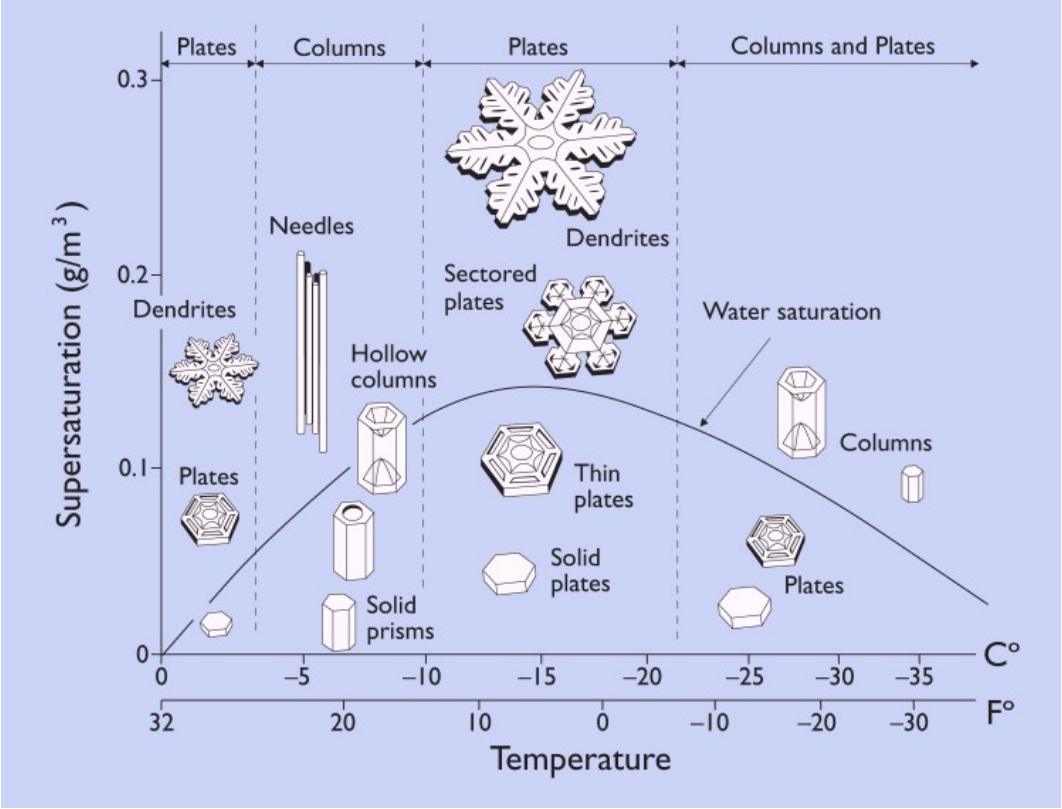


Figure 1.1: Morphology diagram of ice illustrating the shape of snow crystals formed in air at atmospheric pressure, as a function of temperature and water vapour pressure [11].

include an understanding of why growth occurs preferentially along the basal and prismatic surfaces at different temperatures. Fully explaining the origin of the two transitions along the morphology diagram requires an understanding of the growth mechanisms at the crystal surfaces. A detailed picture is required which involves; the crystal structure and its equilibrium surface structure, the kinetics of adsorption and diffusion across the surface, the accommodation of molecules into the crystal lattice, and the vapour density and temperature gradients surrounding the growing crystal.

It is challenging to study ice growth in vapour experimentally due to the need to simplify the parameters and separate the complex effects whist still accurately replicating the growth conditions in a cloud. The surface energies of the crystallographic surfaces are similar and very sensitive to temperature [12, 13]. Accurately determining the conditions of growth within an experiment is itself a major challenge. The experimental ventilation velocity and the presence of water droplets hinders the determination of the true supersaturation environment and prevents comparisons

between the results from different experimental setups [14]. Crystal growth lowers the ambient supersaturation as the vapour supply becomes depleted. The addition of water droplets is therefore required to maintain the water saturation in experiments. However, the presence of the droplets results in an 'effective' supersaturation which is not only difficult to determine but difficult to control. Computer simulations may overcome these challenges and allow us to reduce the complexity of the problem whilst still producing meaningful results and predictions.

In this project we investigate the basal and secondary prism surfaces of hexagonal ice at a range of temperatures between 240 K and 270 K using molecular dynamics and enhanced sampling methods. The secondary prism surface is fast growing and drives growth in lateral direction, whereas the basal surface dominates growth in the vertical direction. Our computer simulations provide insights into the properties of the two surfaces at the molecular level which we correlate with temperature. Our findings help explain the emergence of ice crystals with a specific aspect ratio and the two transitions in the morphology diagram.

1.2 Hexagonal Ice Structure

Ice can be found in at least 13 different crystalline phases depending on the pressure and temperature [1], as shown in the phase diagram in Figure 1.2. Hexagonal ice (I_h) is the most common phase of ice on Earth and can be obtained by freezing water at atmospheric pressure [1]. The oxygen atoms in I_h sit on a regular crystal lattice as shown in Figure 1.3 [1]. Each O atom is surrounded by four oxygen neighbours in a tetrahedral environment. The protons obey the Bernal-Fowler rules which state that [16]:

1. The two hydrogen atoms in each water molecule are directed towards two of the four surrounding oxygen atoms
2. Only one hydrogen atom lies between neighbouring oxygen atoms
3. Each oxygen atom has two nearest neighbouring hydrogen atoms such that the molecular structure of water is preserved

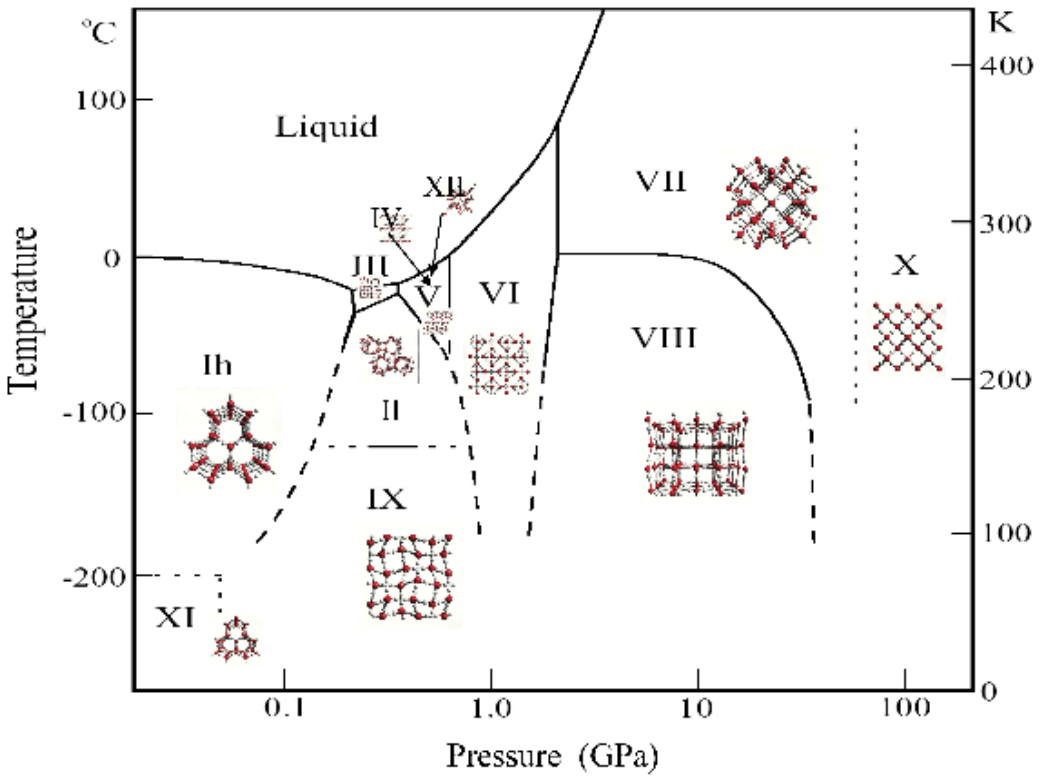


Figure 1.2: Phase Diagram of Ice [15].

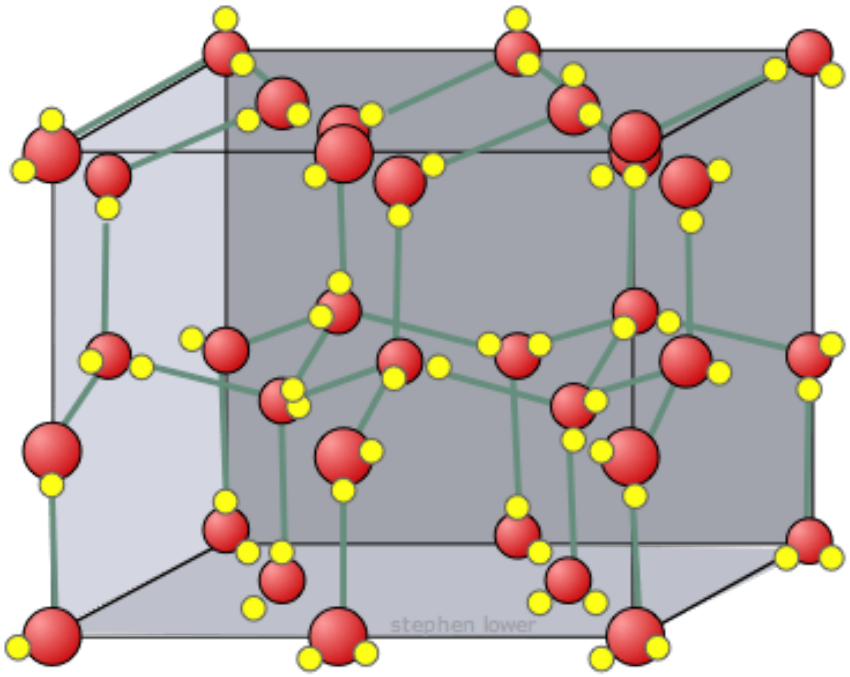


Figure 1.3: Crystal structure of hexagonal ice. Oxygen atoms are illustrated in red and hydrogen atoms in yellow [17].

The crystal structure of I_h is unusual because it contains orientational or proton disorder. Above 5 K, the protons can move between water molecules whilst obeying the Bernal-Fowler rules [18]. The result is electrical polarisability and conductivity [1].

The crystal structure of hexagonal ice is shown in Figure 1.4. I_h forms a low-density structure with a low packing efficiency of $\sim \frac{1}{3}$, calculated as the ratio of the volume of the particles within a unit cell to the volume of the unit cell. In contrast, simple cubic and face centered cubic structures have packing efficiencies of $\sim \frac{1}{2}$ and $\frac{3}{4}$, respectively [19]. Every molecule in hexagonal ice experiences an identical molecular environment. However, there is a small deviation from ideal hexagonal symmetry as the unit cell is 0.3% shorter in the c-direction [19]. The O–O distance is approximately 2.76 Å and each hydrogen atom lies about 0.985 Å from an oxygen atom at around $-20°$C [1].

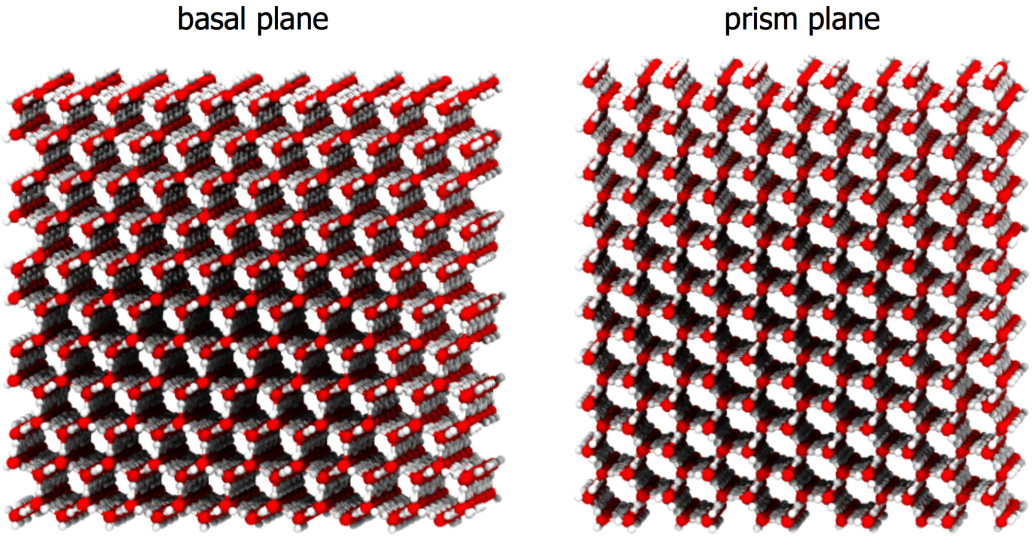


Figure 1.4: Basal and secondary prism planes of hexagonal ice.

Striking images taken by Libbrecht [20] of I_h crystals grown at low vapour pressures are shown in Figure 1.5. The six side faces are the $\{10\bar{1}0\}$ prism faces, and the top and bottom faces are the $\{0001\}$ basal faces, as illustrated in the schematic in Figure 1.5 d).

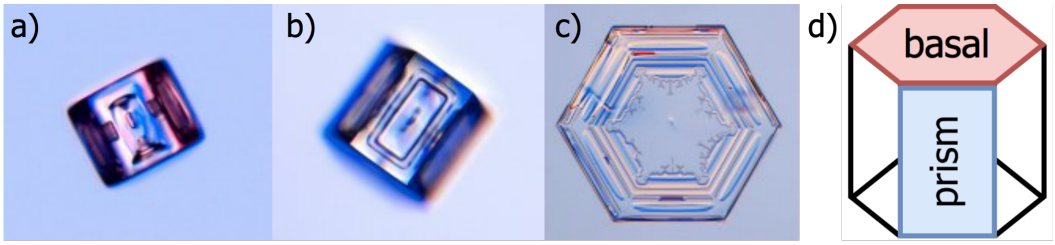


Figure 1.5: Images and schematic of hexagonal ice crystals. a) and b) show elongated prisms and c) shows a thin plate [20]. d) Schematic representation of the basal and prism surfaces.

1.3 Ice Growth - Experiments and Theory

Ice growth in vapour is complex and influenced by numerous effects which are poorly understood. Here we review past works, experiments and theory which seek to explain the influence of thermal diffusion, organic vapours and surface kinetics on the growth of ice in vapour.

Ice growth on a substrate of high thermal conductivity enables growth under conditions of efficient thermal diffusion to be studied. Experiments performed on a stainless steel substrate result in crystallisation which depends on surface kinetic effects more strongly than normal [21]. As the crystal thickness increases, heat transfer to the stainless steel substrate decreases since ice is a poor heat conductor. The ice surface temperature increases due to the latent heat released during crystallisation, resulting in a decrease in the growth rate. Under such conditions, the resistance to heat transfer is the rate limiting factor.

Growth of ice in the presence of organic vapours, such as acetone, acetic acid, nitric acid, silicones and alcohols, can result in significantly different growth shapes at both high and very low vapour pressures [14]. This observation suggests that organic vapour molecules interfere with the ice surface kinetics as opposed to mass and heat transfer processes. Polar foreign vapour molecules modify the growth habit most strongly, suggesting that the kinetic processes modified by the interactions between the polar ice surface and vapour molecules are rate limiting [14].

Many models have been developed to describe the rate of growth of ice in vapour in terms of experimentally measurable quantifies [14, 22]. One such model by Mason [22] quantifies the rate, using the vapour diffusion/thermal conduction equilibrium,

the vapour diffusion/thermal conduction equations and the water vapour density relations, as

$$\frac{dM}{dt} = \frac{4\pi C\sigma}{A+B} \tag{1.1}$$

Mason's model exploits an analogy between the vapour field around an ice crystal and the field of electrostatic potential around a charged conductor of the same size and shape. The leakage of charge from the conductor (the analog of the flux of vapour to or from an ice crystal) is proportional to the electrostatic capacity C of the conductor. C therefore depends on the size and shape of the crystal and accounts for the vapour field. For a sphere of radius r, $C = r$ whilst for a long, thin prolate spheroid $C = \frac{a}{\ln(2l_a/l_b)}$ where l_a is half the length and l_b is the radius of the midsection [14, 23]. A in Eq. 1.1 is a heat term which quantifies the conduction of heat from the growing surface, B is a moisture term which takes into account the transfer of mass to the ice surface, t is time and M is mass. $\sigma = \left[\frac{\rho}{\rho_s(T)} - 1\right]$ is a measure of the supersaturation of the environment, ρ, relative to ice. In this construct, $\rho_s(T)$ the saturation vapour pressure over a plane of ice at temperature T. A is defined as

$$A = \frac{\Delta H_{sub}}{KT}\left[\frac{\Delta H_{sub} M_m}{RT} - 1\right] \tag{1.2}$$

and B is defined as

$$B = \frac{RT}{DM_m\rho_s(T)} \tag{1.3}$$

where M_m is the molecular weight of ice, R is the gas constant, K is the thermal conductivity of air, T is the temperature of the environment, ΔH_{sub} is the latent heat of sublimation, and D is the diffusion coefficient of water vapour in air. At constant pressure, A and B depend only on temperature.

Eq. 1.1 yields the mass growth rate as a function of temperature. This is calculated at two different ambient pressures in [24] using experimental values for the variables and an empirical value for C, and the result is shown in Figure 1.6. The maxima in the trend is a result of the difference between the saturated vapour

pressures over water and ice. As pressure decreases, water vapour diffusivity increases and the rate of growth increases.

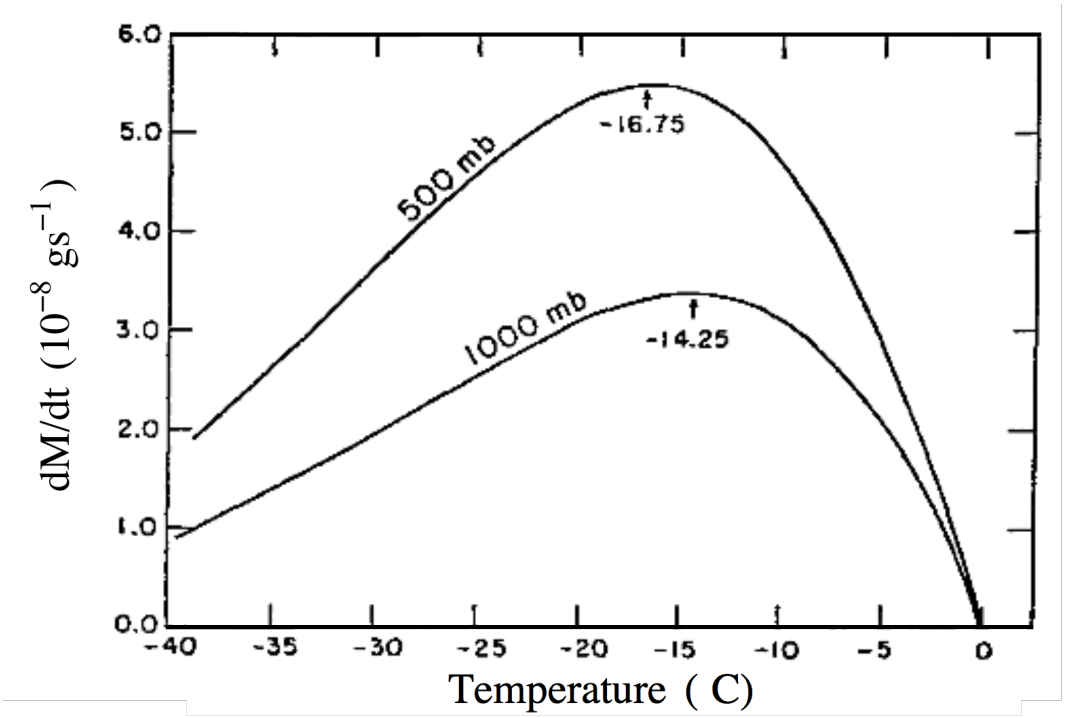


Figure 1.6: Theoretical growth rates of ice crystals as a function of temperature at two water vapour pressures. The temperature of the maximum rate is indicated on each curve [24].

Ice growth modelled using Eq. 1.1 does not take into account surface kinetic effects and assumes that vapour molecules are incorporated directly into the crystal lattice. Any agreement with experimental results would suggest surface kinetic effects are unimportant. However, experimental results do not agree with the simple functional form in Figure 1.6 and surface kinetics must not be overlooked [14]. Eq. 1.1 gives a theoretical prediction of the growth rate of ice crystals in vapour, however, it does not predict how the adsorbing molecules are distributed over the crystal surface or the growth habit, and does not explain the anisotropy of the basal and prism growth rates with temperature.

The influence of surface kinetic effects on the rate of growth of the basal and prism surfaces is studied by experiments performed in pure water vapour. Under such conditions, the growth rate is not governed by the diffusion of water molecules to the crystal surface, enabling the contributions of kinetics to be measured [25]. Ex-

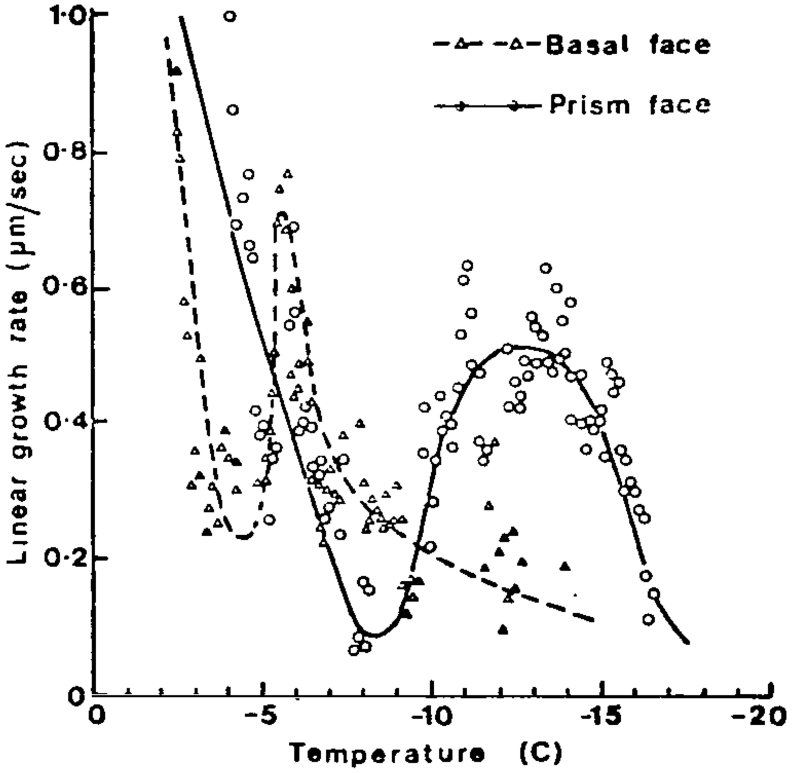


Figure 1.7: Experimental linear growth rate of the basal and prism surfaces of ice as a function of temperature [25].

perimental results suggest that surface kinetics play an important role, and influence the basal and prism surface growth with a magnitude dependent on temperature and crystallographic face [14, 25]. The linear growth rate measured from one such study at constant vapour density and varying temperature is shown in Figure 1.7.

A detailed understanding of the surface kinetics responsible for the curve in Figure 1.7 remains incomplete, however, a quasi-liquid layer on the surface of ice with a varying thickness may account for the temperature dependence of the surface kinetic effects.

1.4 Quasi-Liquid-Layer of Ice

Surface melting or premelting is a phenomenon which occurs on the crystal surface of a range of materials such as metals, semiconductors, and inorganic and organic materials, at temperatures below the melting point [26]. A less ordered quasi-liquid layer (QLL) forms on the surface of these materials, driven by the lowering of the

surface free energy [27]. QLLs are important because they dominate the surface properties of crystals at temperatures below the melting point [28]. As an everyday example, the slipperiness of a skating rink is a product of the QLL of ice. In this project the QLL is of interest because it plays an important role in crystallisation [28]. A water vapour molecule may adsorb on an ice surface from which it will either desorb or become incorporated into the crystal structure. This process is governed by the interaction of the QLL with the vapour molecule. Molecular dynamics (MD) simulations suggest a correlation between the likelihood that a QLL molecule will desorb and its position within the QLL [29]. Ice growth in vapour is therefore mediated by the QLL, and an understanding of the energetics of ice crystal growth requires an understanding of the QLL at the molecular level.

1.4.1 Observing the Quasi-Liquid-Layer

The QLL of ice can be observed experimentally using a range of techniques. Optical ellipsometry and microscopy experiments rely on the different refractive indexes of water and ice, and provide evidence of the QLL on the surfaces of ice [30, 31]. X-ray diffraction experiments reveal that the long-range order, which gives rise to Bragg reflections in the bulk, is lost in a layer close to the surface at all temperatures above $-13\,^{\circ}\mathrm{C}$ [32]. By varying the angle of the incident X-rays, the QLL thickness is measured as a function of temperature above 260 K for the three surfaces of hexagonal ice. The results fit the equation:

$$L = I \ln \left| \frac{T_s}{(T_m - T)} \right| \tag{1.4}$$

where L is the thickness, T_m is the melting temperature, $T_s \sim 260$ K, and $I \sim 8.4$ nm or 4.0 nm for the basal and prismatic surfaces, respectively [1].

Studies of the basal surface of ice by proton back-scattering indicate the presence of an amorphous layer above $-60\,^{\circ}\mathrm{C}$ [33]. Anomalously large amplitudes of molecular vibration are observed near the surface leading to the hypothesis that the QLL is due to large-amplitude thermal vibrations of surface oxygen atoms which propagate into the crystal interior because of the directionality of hydrogen bonds [33]. A thickness

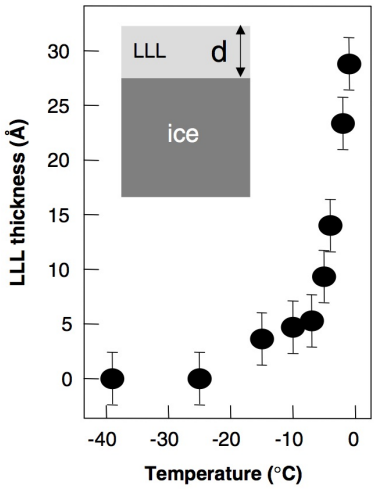


Figure 1.8: Quasi-liquid-layer thickness as a function of temperature measured using photoelectron spectroscopy [34].

of 94 nm is measured at $-1\,^{\circ}\mathrm{C}$ and the functional dependence of the QLL thickness with temperature is found to be:

$$L(\mathrm{nm}) = (94 \pm 17) - (54 \pm 14)\log(273 - T) \tag{1.5}$$

which is in agreement with the functional form of Eq. 1.4.

Photoelectron spectroscopy studies of the surface of ice yield a QLL thickness of about 2 nm at $-2\,^{\circ}\mathrm{C}$, but a negligible thickness below $-20\,^{\circ}\mathrm{C}$ [34]. The temperature dependence of the QLL obtained is shown in Figure 1.8. Atomic force microscopy experiments show a QLL above $-35\,^{\circ}\mathrm{C}$ and a QLL thickness $\propto -\log(T_m - T)$ [35], as observed in X-ray diffraction and proton back-scattering experiments.

In summary, a broad range of experimental techniques reveal the presence of a QLL on the surface of ice. However, the different methods disagree on the temperature at which surface melting first occurs and the QLL thickness. The measured QLL thickness varies by up to 2 orders of magnitude depending on the experimental technique [28]. The source of such significant variations can be attributed to a lack of spatial and temporal resolution and variations in the ice samples [28]. In addition, different techniques use different physical properties of the QLL to measure the thickness, resulting in different sensitivities [27].

1.4.2 Simulating the Quasi-Liquid-Layer

The QLL can be simulated and studied using computer simulations. Selected studies using different water models and ice slab sizes are reviewed in the following.

Classical molecular dynamic simulations are performed by Conde [7] on the vapour exposed basal and prism surfaces of hexagonal ice. Four water models are used: SPC/E, TIP4P, TIP4P/Ice, and TIP4P/2005 which are reviewed in section 3.1. The slabs consist of 1536 water molecules and are $\sim 30 \times 27 \times 60$ Å^3 in dimension. 50 Å of vacuum is used such that the total simulation box is $\sim 30 \times 27 \times 110$ Å^3 in size. NVT simulations are performed for 6 to 12 ns at 200 K, 240 K, 266 K and 270 K. QLL molecules are identified using a tetrahedral order parameter which takes into account the angles between neighbouring oxygen atoms. The results show that the thickness of the QLL depends on the crystallographic face that is exposed but is independent of the water model. QLL formation occurs at about $-100\,^{\circ}$C and $-80\,^{\circ}$C for the basal and prism TIP4P/Ice surfaces, respectively. At 240 K, a QLL thickness of 3.8 Å and 2.7 Å is estimated for the basal and prism TIP4P/Ice surfaces, respectively, whilst at 270 K, the thickness is 7.5 Å and 6.8 Å, respectively. At low temperatures the simulations suggest a QLL is formed of water molecules from just one ice bilayer. This is consistent with more recent MD studies using a variety of water models [36–38].

Neshyba [29] performs 170 ns of molecular dynamics simulations of the basal surface exposed to vacuum at 250 K using a six-site intermolecular potential. The slab consists of 1280 molecules and is $36 \times 31 \times 33$ Å^3 in dimension but is elongated to $36 \times 31 \times 70$ Å^3 to add 37 Å of vacuum. The work reveals that $\sim 92\%$ of the QLL molecules belong to a deeper crystalline sub-layer where each molecule forms typically three hydrogen bonds and has an orientational distribution that closely resembles bulk crystalline ice. The remaining QLL molecules typically have two hydrogen bonds and form a highly mobile outer sub-layer with a preference for dangling bond orientations. Bulk molecules predominantly form four long lived H-bonds which persist for over 1 ns, whereas molecules in the QLL form very few long lived H-bonds. Related work by Pfalzbraff [38] corroborates these findings.

The growth rate of ice in the melt can be estimated from simulations by monitoring the evolution of the length of a growing I_h slab in contact with water [39]. Rozmanov [40] shows that the basal surface growth rate is approximately 30% smaller than the prism growth rate using TIP4P/2005. The position of the interface is defined using a Fourier transform of the density profile. Similar work by Espinosa [39] estimates a maximum rate of growth for the secondary prism surface at 260 K and 230 K using TIP4P/Ice and mW, respectively. The mW maximum rate of growth is three orders of magnitude larger than TIP4P/Ice. For supercoolings of around 55 K, the mW rate is five orders of magnitude larger than TIP4P/Ice. Espinosa defines the interface using the 6^{th} order Steinhardt bond-order parameter described in Section 3.4.1. Both works find that the basal surface growth rate is around 40% slower compared to the secondary prism surface. Below $-15\,^{\circ}\text{C}$, Espinosa [39] describes the dependence of the growth rate, u on supercooling by

$$\ln(u) = \alpha + \beta\Delta T + \gamma(\Delta T)^2 \tag{1.6}$$

where α, β and γ are constants fitted to the simulation data and $\Delta T = T_m - T$ [39]. Rozmanov [40], however, fits the growth rates estimated from simulations of ice in the melt using a slightly different form:

$$u = C_1 \exp\left(C_2/T\right) \sqrt{D(T)}\left[1 - C_3 \exp\left(C_4/T\right)\right] \tag{1.7}$$

where C_1, C_2, C_3 and C_4 are constants fitted to the simulation data, and $D(T)$ temperature dependent self-diffusion constant of the 4-site water model used. Two different equations have been suggested in two similar simulation studies of ice growth in the melt. A lack of conformity of Eq. 1.6 and 1.7 underscores the need for further study of ice growth using simulations and a more fundamental understanding of the growth processes at the molecular level.

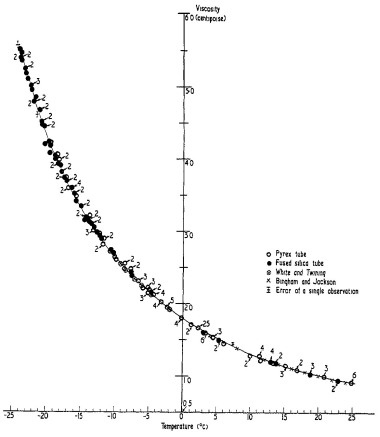


Figure 1.9: Experimental viscosity of water between 25.0 °C and −23.8 °C [41].

1.4.3 Dynamics of the Quasi-Liquid-Layer

The dynamics of the QLL play an important part in the kinetics of ice growth. Studies of the diffusivity and viscosity of ice surfaces have helped uncover the properties of the QLL. Here we review a selection of experimental and simulation works.

The viscosity of supercooled water is measured using a capillary flow technique down to −23.8 °C in Ref. [41]. Figure 1.9 shows the viscosity temperature relationship obtained and includes data from other studies. As temperature decreases, the viscosity increases at an increasing rate.

The temperature dependence of the viscosity of supercooled water is studied by Dehaoui [42] by fitting the experimental results to non-diverging and diverging functional forms. Non-diverging models such as the Arrhenius law, $\eta(T) = \eta_0 \exp(E_a/k_B T)$, and a parabolic law, $\eta(T) = \eta_0 \exp[J^2(1/T - 1/T_0)^2 + (E_a/k_B T)]$, do not reproduce the data correctly. Here η is the viscosity, η_0 is a parameter of the model, J is an energy scale parameter, T_0 is a temperature parameter of the fit and E_a is an energy barrier. Diverging models, in contrast, reproduce the data more accurately. The diverging Vogel-Fulcher-Tamman (VFT) equation, $\eta(T) = \eta_0 \exp[BT_0/(T - T_0)]$ is an improved model compared to the parabolic law, however, it struggles with reproducing the experimental data at high and low temperatures. The power law, $\eta(T) = \eta_0(T/T_s - 1)^{-\gamma}$, where T_s and γ are parameters, overcomes this issue and

gives an excellent fit over the whole temperature interval.

NMR diffusion measurements down to 238 K reveal that in the temperature range from 298 K to 242 K, diffusion is well described by an empirical VFT relationship [43]. Below 242 K, a fractional power-law equation of the form $D = D_0 T^{1/2} \left(\frac{T}{T_s} - 1\right)^{\gamma}$ provides a better fit. NMR studies of ice between 253 and 273 K yield in-plane diffusion coefficients between 1.3×10^{-9} $\mathrm{cm^2 s^{-1}}$ and 3.08×10^{-9} $\mathrm{cm^2 s^{-1}}$ and an activation barrier of 23.2 kJ/mol [44]. Simulations estimates of the in-plane diffusion coefficient within the first bilayer using the rigid 4-site TIP4P water model range from 1.7×10^{-7} $\mathrm{cm^2 s^{-1}}$ at 180 K to 1.6×10^{-6} $\mathrm{cm^2 s^{-1}}$ at 210 K. An average in-plane diffusion energy barrier of 23.2 ± 2.9 kJ/mol is estimated which is close to the energy required to break a hydrogen bond, suggesting that in-plane diffusion occurs via repeated breaking and formation of H-bonds [45].

Detailed calculations reveal that TIP4P/Ice underestimates the diffusion coefficient by a factor of two at moderate supercooling and by a factor of four at high supercoolings [39]. In contrast, the mW diffusion coefficient is up to three orders of magnitude larger than the experimental result and is much less strongly dependent on temperature compared to TIP4P/Ice [39]. Above 230 K, the TIP4P/Ice diffusion fits $D = D_0(T - T_{MCT})^{\alpha}$ whilst at lower temperatures the results are better described by an Arrhenius equation [39]. The change in temperature dependence of the diffusion coefficient is denoted as the fragile-strong transition [46].

Computational studies of the QLL diffusivity provide further insights into the QLL properties, and enable the accuracy of water models to be assessed. The self-diffusivity of the prism surface is studied using molecular dynamics simulations in Ref. [47] with a 4-site, 5-site and 6-site water model. An $\mathrm{I_h}$ system consisting of 2880 water molecules is simulated at submelting temperatures ranging from $-2\,^{\circ}\mathrm{C}$ to $-60\,^{\circ}\mathrm{C}$. Below $-40\,^{\circ}\mathrm{C}$, the surface diffusion is anisotropic, whilst at higher temperatures, above $-30\,^{\circ}\mathrm{C}$, it is isotropic. The results show that anisotropic diffusion occurs when the diffusion activation energy is low and of the order of one hydrogen bond. Under such conditions, the diffusion mechanism is dominated by motion within the outermost QLL sublayer where the entropic barriers to diffusion are different in

the two in-plane directions and the enthalpies of activation are similar in the two in-plane directions. A mechanism for in-plane anisotropic diffusion within the QLL involving uncorrelated hopping or jumping, as opposed to molecular exchanges, is suggested based on MD simulations [47]. Isotropic diffusion, however, occurs when the activation energy is larger and results in a different mechanism within a deeper QLL sublayer. The mechanism involves a more fluid QLL that is less sensitive to the underlying crystalline lattice.

An activation energy equivalent to the energy of one hydrogen bond is estimated for horizontal movement of water molecules within the topmost QLL sub-layer, using the six-site NE6 intermolecular potential in Ref. [38]. Simulations are performed at 250 K for 200 ns on slabs consisting of 2880 water molecules and of dimension $54 \times 47 \times 37$ Å^3. Vacuum is added by elongating the simulation box by 33 or 43 Å, such that the total simulation boxes is either $54 \times 47 \times 80$ Å^3 or $54 \times 80 \times 37$ Å^3. The results show that the surface diffusivity of hexagonal ice is dependent on the crystallographic surface. Surface diffusivity is isotropic on the basal surface but anisotropic on the prismatic surface. Analysis of the vertical coordinates of the oxygen atoms yields insight into the dynamics of layer-to-layer transitions. Sublimation events are found to be rare whilst transitions between the QLL sub-layers are relatively common. Transitions from the outermost ice sub-layer into the bottom QLL sub-layer occurs as binary exchanges [38].

A more detailed MD study of the activation energy of diffusion estimates a diffusion energy barrier equivalent to one hydrogen bond at low temperatures increasing to two hydrogen bond close to the melting point [48]. This result affirms a change in mechanism of diffusion which accompanies the observation of a transition from anisotropic to isotropic in-plane diffusion as temperature increases. In contrast, measurements of the activation energy of diffusion of supercooled water decreases with temperature suggesting a radically different mechanism and implying that the underlying crystalline lattice of ice plays and important role [48].

The dynamics of the QLL are complex and the in-plane diffusion of the two surfaces are isotropic and anisotropic at different temperatures. Diffusion within

the QLL occurs via a mechanism involving an initial migration to the outermost sublayer of the QLL followed by in-plane motion with activation energy equivalent to around one hydrogen bond [37, 38, 45]. The molecular origin of the convoluted temperature dependence of the in-plane diffusion remains unknown, and the impact on the rate of growth unmapped. A detailed understanding of the basal and prism surface structure is therefore necessary for a comprehensive picture of ice growth.

1.5 Surface of Ice - Structure and Energy

The surface structure of ice is complex and poorly understood and yet important for understanding the dynamics of the QLL. Ice crystal growth is a dynamical process which depends on the surface structure and therefore also remains poorly understood. Surface melting occurs over the temperature range where a large variation in snow crystal morphology occurs which suggests that surface melting plays an important role in crystal growth [11]. However, it is not clear how surface melting influences the surface structure or how it affects the attachment kinetics central to crystal growth [49]. The functional form of the QLL thickness with temperature suggests that the QLL diverges to infinite thickness as the melting temperature approaches. As a result, a reasonable assumption is that the QLL/ice interface resembles the liquid/ice interface close to the melting temperature. However, at lower temperatures the QLL structure must be characterised in order to understand the dynamics and energetics of ice crystal growth.

The protons of I_h follow the ice rules and are fully orientationally disordered close to the melting point. I_h remains the stable phase under ambient pressures down to ~ 72 K [50]. Below this temperature, the ferroelectric proton ordered phase, ice XI, becomes the ground state. However, as temperature is decreased, I_h proton motion is reduced and a glassy transition is observed near 110 K [51] which, in practice, prevents transitions to ice XI.

Antarctic ice sheets kept at low temperatures for thousands of years can provide insights into the equilibrium proton structure. Raman and neutron scattering of

Antarctic ice sheets suggest a second-order phase transition to a proton-ordered phase occurs at 237 K [52]. However, more recent studies cast doubts on this discovery and suggest ancient Arctic and Antarctic ice have the same proton disordered structure as ice I_h [53, 54].

Whist the proton ordering of bulk ice I_h has been widely studied [55, 56], the ordering and energetics of I_h surfaces are not yet fully understood [57, 58]. The entropy at the surface is greater than in the bulk and varies depending on the crystallographic surface [59]. The larger the entropy, the more favourable the surface, however, the impact of entropy on the growth mechanisms is unclear. The surface energy affects crystal growth and influences the equilibrium crystal shape and yet the value for ice is not well established both experimentally and theoretically, whilst its dependence on proton ordering is not rigidly known [58]. DFT calculations suggests significant differences between the energetics of proton ordering on the surface and in the bulk, and a much higher order-disorder transition temperature at the surface compared with the bulk [58].

The interface energy between ice and vapour can be measured experimentally from the contact angle of a water drop on an ice surface. However, it is unclear how the droplet influences the surface structure or whether intermediate structures form which influence the experimental measurement [60]. The surface energy of ice measured at $-25\,^{\circ}\mathrm{C}$ is found to be roughly 4.8 meVÅ^{-2} [61]. Estimates obtained using empirical methods, such as electronegativity based models [62], and semi-empirical methods, which combine theory with experimental measurements, range from 4.3 meVÅ^{-2} to 18.6 meVÅ^{-2} [60]. Comparisons of the surface energy estimates suggest they depend strongly on the model and parameters used [60]. *Ab initio* simulations can overcome these issues and provide an estimate of the surface energy of ice.

The basal surface energy estimated using total energy DFT calculations reveal a range of values increasing from 12.2 meVÅ^{-2} to 18.2 meVÅ^{-2}, as the proton order of the configuration is decreased [60]. In contrast, the range of bulk energies obtained as the proton order is varied is an order of magnitude smaller. Similar results are

obtained for the prism surface [60].

The significant difference in dependence of bulk and surface energies on proton arrangement is due to repulsion from dangling OH groups. Dangling OH bonds are defined as OH groups which point out of the surface. At the surface these are charge-charge interactions with a $1/d$ dependence, where d is the atomic separation between two dangling bonds [60]. In the bulk, dangling OH groups interact through weaker dipole-dipole interactions [60]. One hypothesis suggests that surface melting occurs preferentially at regions with higher concentrations of dangling OH groups [58]. Whilst the surface and bulk energies depend to differing extents on the proton order, the O-O bond lengths are very similar. Experimental results at 38 K show that the O-O bond length of surface molecules is 2.77 Å, compared to 2.76 Å in the bulk [63].

The vertical binding energy (VBE) of a water molecule is defined as the difference between the total potential energy of the configuration containing the molecule and an identical configuration with the molecule removed [29, 38]. MD simulations of the basal and prism surfaces at around $-40°$C suggest a VBE equivalent to four hydrogen bonds within the bulk [29, 38]. Within the QLL, a VBE of around two H-bonds is detected in the outer sublayer whilst a VBE of around three H-bonds is measured deeper within the QLL. A broad range of molecular VBEs are estimated within the QLL suggesting that the QLL is characterised by a broad range of hydrogen bonding networks. Analysis of the lifetime of H-bonds suggests that QLL molecules have predominantly short-lived H-bonds whilst bulk molecules have long-lived H-bonds [29].

Vibrational sum frequency generation (SFG) spectra reveal information on the vibrational response of surface molecules at the solid/vapour interface. SFG spectra enable the strength of hydrogen bonds between neighbouring molecules to be quantified based on the measured frequency of OH stretch vibrations [64, 65]. Details of the mechanism of ice growth from the QLL are uncovered by Sanchez [66] by combining spectra calculated from classical molecular dynamics simulations with SFG spectroscopy measurements. The simulations are performed using a 4-site TIP4P/Ice

water model described in section 3.1. The results provide evidence of layer-by-layer growth. SFG spectra of the basal surface reveal a shift at around 257 K indicating an abrupt weakening of hydrogen bonds above 257 K. Correlating the density profiles of the basal surface with the SFG results indicates that the surface melts layer by layer which is in agreement with previous work [67]. The results suggest that at 257 K there is a shift from one melted bilayer to two melted bilayers. Computational studies of the orientation of OH groups support this view and reveal a noticeable shift in trend for temperatures above and below 257 K [66]. The emerging picture of ice surface melting is that the first bilayer begins to melt at 200 K, and at 257 K a second bilayer melts. Detailed DFT calculations show that the molecular binding energy associated with the removal of a water molecule from ice varies by 30 to 70 kJ/mol in the first two bilayers but is uniform in the third [68]. This result indicates a difference in bonding network within the first two bilayers and subsequent layers. The melting behaviour of the first two bilayers could therefore be different to deeper ice layers. The vibrational response of dangling OH bonds show a weak temperature dependence which suggests that the outermost surface structure remains unchanged between 235 K and 264 K [66]. This result is consistent with previous findings [69].

At higher temperatures a different picture emerges from high-resolution optical interferometry measurements [70, 71]. Close to the melting temperature, the QLL is composed of a thin liquid layer. Sitting on top of this layer are micrometer sized water droplets which come in and out of existence and are dependent on the water vapour pressure. The results paint a complex picture of a spatially and temporally heterogeneous QLL. Droplets form due to long-ranged attractive interactions between ice and vapour. The QLL is suppressed by these interactions and the balance between QLL formation and droplets formation is determined by the competition between the long-ranged attractive interactions and short-ranged repulsions which enhance the QLL [72]. The result of these competing effects is a first-order transition between different wetting states [66]. A droplet of liquid may form directly on the surface of ice or directly on top of a QLL. The former is favoured by the long-range interactions whilst the latter is favoured by short-ranged forces.

Short-ranged repulsive interactions change the temperature dependence of the QLL thickness from a power law predicted for dispersion forces to a logarithmic divergence [72].

The overarching consensus from a broad range of studies of the surface of ice is that it is complex and depends on both temperature and the crystallographic plane. Building a complete picture of ice growth in vapour is very challenging because each individual step in the crystallisation process depend non-trivially on a large number of factors. Previous works, described in detail in the previous sections, have focused on probing individual steps in detail, however, combining all the different factors and contributions into a global picture remains a significant challenge for the community.

1.6 Aim of this project

The shape of vapour grown ice crystals depends strongly on the conditions under which growth has taken place. As mentioned in section 1.1, at low vapour pressures ice growth results in the formation of either elongated hexagonal prisms or thin hexagonal plates depending on the temperature. The shape depends on the relative rate of growth of the exposed basal and prismatic faces. Elongated prisms occur when the rate of growth of the basal face is fast compared to the rate of growth of the prismatic faces, whilst thin plates occur when the opposite is true. As temperature increases, there are two cross-overs in the rates of growth of the two surfaces. However, the reasons why growth occurs preferentially along the basal and prismatic surfaces at different temperatures remain unknown. The underlying mechanisms at each step in the cycle of a crystallising vapour molecule are not fully understood, hindering progress in climate science.

Whilst theoretical models have been developed to describe the growth rate of ice in vapour such as Eq. 1.1 described in section 1.3, there is no explanation why at certain temperatures the basal surface grows fastest and at others the prism surface grows fastest. A complete understanding of the rate of growth of the basal and prismatic surfaces requires not only an understanding of vapour adsorption but also

a detailed knowledge of the processes which occur once a molecule has absorbed.

In this work we study the growth of ice in vapour using computer simulations. We focus on both the QLL/vapour and QLL/ice interface dynamics and uncover details of the mechanisms of growth. We build a comprehensive description of the basal and prismatic surfaces involving the crystal and equilibrium surface structures, the dynamics of the QLL, the thermodynamics of QLL crystallisation and melting, the energetics at the QLL/vapour interface, and the dynamics at the QLL/ice interfaces. We focus on the secondary prism surface of ice as opposed to the primary prism surface because it is the fastest growing prismatic surface and drives lateral growth. In this thesis the secondary prism surface is refered to as simply the prism surface when discussing our simulation results. By combining the different elements of our work, we aim to explain the origin of the two cross-overs in the morphology diagram of ice.

Molecular dynamic simulations are used in this project to study the QLL and its properties using two water models. Diffusion related properties are calculated and structural analysis unearths important differences between the basal and prism surfaces. Our molecular dynamic simulations provide reliable statistics on the crystal surface dynamics of ice as we model larger systems than previous works. The free energy barriers associated with QLL melting and crystallising are recovered using an enhanced sampling technique (Metadynamics) and provide a measure of the thermodynamic chemical potential driving force of QLL crystallisation.

Enhanced sampling techniques have never been used before to study the role of the QLL of ice in crystallisation. In this project we obtain the first estimates of the free energy associated with growth of the basal and prismatic surfaces of ice. Our focused approach reveals important insights into the growth of ice which contribute towards improved climate modelling and can help tackle the impact of global warming. The scope of our work, however, is not just restricted to studies of ice. Solids interact through their surfaces and their shape will influence their chemical and physical properties. Drugs are often purified and sold as crystalline solids, therefore their size and shape will not only have an impact on the efficiency of the production

processes, but will influence the drugs dissolution kinetics, bio-availability and uptake performances [73]. The approach used in this project is therefore also valuable to the fields of pharmacology and fine chemistry, due to an interest in the ability to control the shape of crystals [73].

Chapter 2

Simulation Methods

The growth of ice is studied in this thesis using computer simulations. The following chapter reviews the simulation methods used and outlines the algorithms.

2.1 Molecular Dynamics

Newton's equation of motion, $\mathbf{F} = M\mathbf{a}$, suggests it is possible to determine the evolution of the position of a particle with time, by calculating its acceleration, $\mathbf{a}$ from the force it experiences, $\mathbf{F}$, and its mass, M. Based on this notion, classical molecular dynamics (MD) simulations allow the equilibrium and transport properties of many-body systems to be computed by treating the particles using classical mechanics. In atomistic MD simulations, the particles correspond to atoms. In coarse-grained simulations, groups of atoms are represented by one particle.

2.1.1 Born-Oppenheimer approximation

The evolution of a many-body system is in principle described by solving the time-dependent Schrodinger equation

$$H(\mathbf{R}, \mathbf{r})\Psi(\mathbf{R}, \mathbf{r}) = E\Psi(\mathbf{R}, \mathbf{r}) \tag{2.1}$$

where $\mathbf{r}$ and $\mathbf{R}$ are the set of electronic and nuclear coordinates, respectively. The Hamiltonian operator can be broken down into 5 main contributions

$$H(\mathbf{R},\mathbf{r}) = T_n(\mathbf{R}) + V_{nn}(\mathbf{R}) + T_e(\mathbf{r}) + V_{ee}(\mathbf{r}) + V_{en}(\mathbf{R},\mathbf{r}) \tag{2.2}$$

where T_n and T_e correspond to the nuclear and electronic kinetic energy operators, respectively, and V_{nn}, V_{ee} and V_{ne} correspond to the nuclear-nuclear, electron-electron and nuclear-electron interaction potential energy operators, respectively.

An exact solution to Eq. 2.1 is not possible for an N-body system, therefore approximations must be made. Electrons are three orders of magnitude lighter than nuclei and yet their electric charge is of the same order. The result is that both nuclei and electrons experience similar forces, however, electrons have much greater velocities.

The Born-Oppenheimer approximation is based on the assumption that within the time-scale of nuclear motion, electrons can relax to the instantaneous ground-state configuration. Nuclei are considered stationary, allowing the energy of the system to be calculated based on the electronic ground-sate configuration, and allowing the nuclear motion to be solved. The Born-Oppenheimer approximation leads to a separation of electronic and nuclear motion, allowing the wavefunction to be expressed as

$$\Psi(\mathbf{R},\mathbf{r}) = \chi(\mathbf{R})\phi(\mathbf{r};\mathbf{R}) \tag{2.3}$$

where $\phi(\mathbf{r};\mathbf{R})$ is an electronic wave function parameterised by the nuclear positions, $\mathbf{R}$ and $\chi(\mathbf{R})$ is the nuclear wave function. Following from this, we may split the Schrodinger equation into an electronic contribution

$$H_e(\mathbf{r},\mathbf{R})\phi(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R})\phi(\mathbf{r};\mathbf{R}) \tag{2.4}$$

and nuclear motion contribution

$$\left[T_n(\mathbf{R}) + E_e(\mathbf{R})\right]\chi(\mathbf{R}) = E\chi(\mathbf{R}) \tag{2.5}$$

where $E_e(\mathbf{R})$ is the potential energy for nuclear motion and the electronic Hamiltonian

is defined as

$$H_e(\mathbf{r}, \mathbf{R}) = V_{nn}(\mathbf{R}) + T_e(\mathbf{r}) + V_{ee}(\mathbf{r}) + V_{en}(\mathbf{R}, \mathbf{r}) \tag{2.6}$$

Molecular dynamics simulations usually treat nuclei as classical particles and assumes that the quantum effects on nuclear motion is negligible due to their size. As such, the Schrodinger equation of nuclear motion can be substituted by Newton's equation of motion

$$-\nabla_{\mathbf{R}} E_e = M\ddot{\mathbf{R}} \tag{2.7}$$

The evolution of a system with time is obtained from a molecular dynamics simulation by computing the forces the nuclei experience from the negative of the potential energy gradient and integrating the equation of motion Eq. 2.7.

In *ab initio* molecular dynamics, E_e is obtained by solving Eq. 2.6. In classical molecular dynamics simulation, the nuclei follow classical mechanics and move within an effective potential, E_e produced by the electrons. E_e depends only on the position of the nuclei at time t and is obtained empirically.

2.1.2 The Verlet Algorithm

A molecular dynamics simulation involves numerically solving Newton's equation of motion and adjusting the particle positions accordingly at discrete time intervals. The time interval is referred to as the time step, δt. Smaller values of δt result in more accurate simulations. To perform MD simulations, we must be able to calculate the forces that each particle experiences and have a method to integrate the equations of motion.

The Verlet algorithm is widely used to numerically integrate Newton's equation of motion due to its simplicity and stability. The algorithm is exactly time reversible. The derivation begins with a Taylor expansion of the particle position vector with time

$$\begin{aligned}\mathbf{R}(t+\delta t) &= \mathbf{R}(t) + \frac{d\mathbf{R}(t)}{dt}\delta t + \frac{d^2\mathbf{R}(t)}{dt^2}\frac{\delta t^2}{2} + \frac{d^3\mathbf{R}(t)}{dt^3}\frac{\delta t^3}{6} + O(dt^4) \\ &= \mathbf{R}(t) + \mathbf{v}(t)\delta t + \frac{\mathbf{F}(t)}{M}\frac{\delta t^2}{2} + \frac{d^3\mathbf{R}(t)}{dt^3}\frac{\delta t^3}{6} + O(dt^4)\end{aligned} \tag{2.8}$$

where $\mathbf{F}$(t) is the force, M is the particle mass and $\mathbf{v}(t)$ is the velocity. Similarly,

$$\mathbf{R}(t-\delta t) = \mathbf{R}(t) - \mathbf{v}(t)\delta t + \frac{\mathbf{F}(t)}{M}\frac{\delta t^2}{2} - \frac{d^3\mathbf{R}(t)}{dt^3}\frac{\delta t^3}{6} + O(dt^4) \tag{2.9}$$

Combining Eq. 2.8 and Eq. 2.9 results in

$$\mathbf{R}(t+\delta t) = 2\mathbf{R}(t) - \mathbf{R}(t-\delta t) + \frac{\mathbf{F}(t)}{M}\frac{\delta t^2}{2} + O(dt^4) \tag{2.10}$$

which is the essence of the Verlet algorithm. A system is propagated forward by a time step δt using only the coordinates from the previous two time steps and the forces at the current time step. The truncation error of the algorithm is of the order of $O(dt^4)$. The velocities can be obtained from

$$\mathbf{v}(t) = \frac{\mathbf{R}(t+\delta t) - \mathbf{R}(t-\delta t)}{2\delta t} + O(dt^2) \tag{2.11}$$

and have an error of the order of $O(dt^2)$. The Verlet algorithm has a number of disadvantages. It requires that two sets of position vectors be stored in memory and is not self-starting as it relies on the knowledge of the positions at the two previous time steps. In addition, the velocity at time t can only be computed at time $t + \delta t$.

The velocity Verlet algorithm is a reformulation of the Verlet algorithm with an improved accuracy. It is obtained by manipulating the Taylor expansions and can be expressed as

$$\mathbf{R}(t+\delta t) = \mathbf{R}(t) + \mathbf{v}(t)\delta t + \frac{\mathbf{F}(t)}{M}\frac{\delta t^2}{2} \tag{2.12}$$

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t) + \frac{\mathbf{F}(t+\delta t) + \mathbf{F}(t)}{M}\frac{\delta t^2}{2} \tag{2.13}$$

The initial velocities are typically drawn randomly from a Maxwell-Boltzmann distribution at the desired temperature. Using this approach, however, can lead to non-zero net momentum which would result in translational drift of the system during a simulation. This is overcome by subtracting the net momentum from the initial velocity of each particle.

2.1.3 Force Fields

Traditional molecular dynamics simulations use an empirical force field, U, to determine the forces. U is an empirical approximation of the potential energy for nuclear motion, $E_e(\mathbf{R})$.

The force field describes the dependence of the energy of a system on the particle coordinates. It is a simplified model of the true potential that can be evaluated quickly but is sufficiently detailed in the region being simulated to reproduce the properties of interest. The parameters of the force field are typically obtained either by fitting to experimental data or from *ab initio* or quantum mechanical calculations. Bonds cannot be broken or formed using traditional force fields, however, this limitation can be overcome using bond-order based force fields such as ReaxFF [74].

The force field can be split into bonded and nonbonded contributions.

$$U = U_{bonded} + U_{nonbonded} \tag{2.14}$$

Bonded contributions are intramolecular two-, three- and four-body interactions. These include a two-body bond potential which describes the vibrations of covalently bonded atoms, a three-body bond angle potential which describes the angular vibrational motion of bonded atoms, and a four-body torsion angle potential which describes the rotation around a bond. The bond and bond angle potentials describe the energetic cost of deforming the bond length and bond angles from their respective equilibrium values.

$$U_{bonded} = \sum_{bonds} U_{bond} + \sum_{angles} U_{angle} + \sum_{torsions} U_{tors} \tag{2.15}$$

Intermolecular interactions are contained within the nonbonded contributions. These interactions involve all pairs of atoms but usually excludes pairs already involved in a bonded potential contribution. Nonbonded terms consist of Van der Waals and electrostatic interactions:

$$U_{nonbonded} = \sum_{van\ der\ Waals} U_{vdw} + \sum_{electrostatic} U_{el} \tag{2.16}$$

Computing $U_{nonbonded}$ is the most time consuming step in a molecular dynamics simulation. Van der Waals Forces are commonly described by the Lennard-Jones potential

$$U_{vdw} = \epsilon \left[\left(\frac{R_m}{R} \right)^{12} - 2 \left(\frac{R_m}{R} \right)^{6} \right] \tag{2.17}$$

where ϵ is the depth of the potential well, R is the atomic separation and R_m is the atomic separation corresponding to the minima of the potential. The repulsive term behaves as R^{-12} and originates from the overlap of the electron clouds of two atoms. The attractive term behaves as R^{-6} and encapsulates the Keesom force between two permanent dipoles, the Debye force between a permanent dipole and induced dipole and the London dispersion force between two instantaneously induced dipoles. Although Van der Waals interactions are weak compared to electrostatic interactions or covalent bonds, they are hugely important in water.

Electrostatic interactions between two atoms are described by the Coulomb potential

$$U_{el} = \frac{q_1 q_2}{4\pi\epsilon_0 R} \tag{2.18}$$

where q_1 and q_2 is the charge of the two atoms and ϵ_0 is the permittivity of space. Determining the atomic charges is often non trivial. Quantum mechanical calculations can yield the molecular electron density to a high accuracy, however, obtaining a reduced description that can be used in molecular dynamics simulations is challenging.

The most common approach used to calculate reliable partial atomic charges involves computing the quantum mechanical potential from *ab initio* calculations.

Direct assessment of the derived partial charges is not possible since atomic charges are not experimental observables.

Fixed atomic partial charges cannot fully describe polarisation effects in the condensed phase. These effects are usually taken into account by using effective charges. The resulting dipole moment of such models is often larger than the experimental value. One limitation of such force fields is the reduced transferability across different phases. For example, the partial charges used to describe the condensed phase will not be adequate for describing the gas phase. Polarisable force fields are able to overcome the limitations of non-polarisable models but are more computationally costly.

The force acting on the i^{th} particle is determined by the gradient of the potential with respect to particle displacements:

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} U(\mathbf{R}_i, ..., \mathbf{R}_N) = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}\right) \tag{2.19}$$

2.1.4 Boundary Conditions

Periodic boundary conditions (PBC) are used in molecular dynamics simulations. A particle which leaves the cell reappears on the opposite side and thus the simulation box is surrounded by an infinite number of replicas of itself. However, this approach introduces an artificial periodicity to the system which presents issues when simulating a disordered system. Simulating a sufficiently large system can overcome this problem and minimise the effect. Simulations of macromolecules in solution must have a large enough cell to prevent the macromolecule from interacting with its own image across the boundaries and care must be taken when considering properties which depend on long-range correlations.

Computational efficiency can be improved using a spherical truncation which limits interactions to pairs of particles separated by a distance of less than R_c

$$U_{trunc}(R) = \begin{cases} U(R), & R \leq R_c, \\ 0, & R > R_c \end{cases} \tag{2.20}$$

R_c must be less than half of the width of the cell in each direction to avoid interactions with the periodic images but must also be large enough to capture the important features of the potential. The interplay between the minimum and maximum size of R_c prevents the truncation approach to be applied to the electrostatic potential as it does not decay with distance rapidly enough. Instead the Ewald summation method is used [75].

The Ewald sum splits the electrostatic potential into two components. One which decays rapidly with distance and one which contains the long-range interactions and has weak contributions at small distances. The Coulombic potential for ions under PBC can be expressed as

$$U_{el} = \frac{1}{4\pi\epsilon_0 r} \sum_{n} \sum_{i,j} \frac{q_i q_j}{|\mathbf{R}_{ij} + \mathbf{n}L|} \tag{2.21}$$

which includes interactions between periodic images. This potential converges very slowly due to the infinite sum in $\mathbf{n}$ and the result is conditionally convergent because it depends on the order of the summation. The Ewald sum converges rapidly and to an absolute value. The derivation involves defining the potential field generated by an ion as

$$\phi_i(\mathbf{R}) = \frac{1}{4\pi\epsilon_0} \frac{q_i}{|\mathbf{R} - \mathbf{R}_i|} \tag{2.22}$$

where q_i is the charge and $\mathbf{R}_i$ the position. The potential field generated by all ions in the system including their periodic images but excluding ion i can be split into a short range and long range contribution

$$\phi_{[i]}(\mathbf{R}) = \phi^S_{[i]}(\mathbf{R}) + \phi^L_{[i]}(\mathbf{R}) \tag{2.23}$$

leading to the splitting of the Coulomb interaction energy. Following from this it can be shown that the short range interaction energy can be truncated and computed in real space. The long range interaction cannot be directly computed in real space and is transformed into a sum in reciprocal space by applying a Fourier transform to

$\phi^L(\mathbf{R})$. The summation of the long -range interactions is short-ranged in reciprocal space and can be truncated.

Conceptually the Ewald sum involves adding a charge distribution of opposite sign to each point charge. The result is a screened charge which cancels out at long distances and therefore is short-ranged and can be treated in real-space. A compensating Gaussian distribution is added to take into account the added charge distribution which behaves like point charges at large distances.

The computational cost of the Ewald sum scales as $\mathcal{O}(N^{3/2})$. The more recent Particle Mesh Ewald (PME) method [76] scales as $\mathcal{O}(N\log(N))$ and is more efficient than the Ewald sum when N is large.

2.1.5 Constraints

In a molecular dynamics simulation, δt must be smaller than the fastest motion in the system to ensure accurate and reliable simulations. It is, however, desirable to have the largest possible δt to increase the efficiency of a simulation. In atomic systems, the fastest motion corresponds to intramolecular bond vibrations. Freezing these high frequency vibrations by applying constraints to the bond angles and lengths enables a larger time step to be used. These constraints can be applied using the SHAKE algorithm [77] which fixes the bond lengths to a constant value, l_0. In the first step of the SHAKE algorithm, atoms are displaced based on the Verlet algorithm. The resulting atomic separations at this stage do not satisfy the SHAKE distance constraint. In the second step, the deviation in bond length from l_0 is used to calculate a constraint force which retrospectively restores the bond length to l_0. Step two is repeated until every bond length satisfies the SHAKE condition within a tolerance.

2.1.6 Thermodynamics ensembles

Simulations performed using the integration Eq. 2.19 are in the microcanonical ensemble (NVE) and maintain a constant number of particles N, volume V and total energy E. Experiments are usually performed at constant temperature or constant

pressure, therefore, it is desirable to have a way of controlling the temperature and pressure during a simulation. This can be achieved using a thermostat and barostat which modify the equations of motion.

Thermostats

The temperature of a system is correlated to the kinetic energy and can therefore be controlled by modifying the particle velocities. The simplest approach to control temperature is by scaling the particle velocities by a factor $\sqrt{T_0/T(t)}$ where T_0 is the desired temperature and T is the current simulation temperature. However, this scaling approach does not result in a true NVT ensemble and can lead to unphysical dynamics.

A more sophisticated and accurate approach to control the temperature is to couple the system with a heat bath. The Berendsen thermostat [78] uses this approach and controls the temperature by rescaling the velocities at regular intervals by the following factor

$$\sqrt{1 - \frac{\delta t}{\tau}\left(1 - \frac{T_0}{T}\right)} \tag{2.24}$$

where δt is the time step and τ is a constant that determines the strength of the coupling between the heat bath and the system. A large τ corresponds to weak coupling and a smooth and slow update of the temperature during the simulation. When τ is equal to the time step, the Berendsen thermostat is equivalent to the simple velocity scaling method. When $\tau \to \infty$, we return to the NVE ensemble. The Berendsen thermostat and velocity scaling method suppress fluctuations in the kinetic energy and therefore do not generate a true canonical ensemble.

The Andersen thermostat [79] couples the system to a heat bath and is able to generate the correct canonical ensemble. The coupling between the bath and system manifests itself through stochastic forces that act occasionally on randomly selected particles. The thermostat mimics collisions between system particles and bath particles at a specified temperature. The collision frequency determines the coupling strength. A particle involved with a stochastic collision receives a new

velocity selected from a Maxwell-Boltzmann distribution at the bath temperature T_0. Assigning particles random velocities alters the dynamics, resulting in unphysical diffusivities and viscosities which depend on the collision frequency [80].

The Nosé-Hoover thermostat treats the heat bath as an integral part of the system. It achieves this by reformulating the equations of motion using the Lagrangian approach to classical mechanics.

The Lagrangian Method

The Lagrangian, L, of a system is a function of the coordinates and velocities and is defined as the difference between the kinetic and potential energies, T and U respectively.

$$L(\{\mathbf{R_i}\},\{\dot{\mathbf{R}}_\mathbf{i}\}) = T(\{\dot{\mathbf{R}}_\mathbf{i}\}) - U(\{\mathbf{R_i}\}) = \frac{1}{2}\sum_i m_i \dot{\mathbf{R}}_\mathbf{i}^2 - U(\{\mathbf{R_i}\}) \tag{2.25}$$

We can define an action, S, as the integral of the Lagrangian between t_i and t_f

$$S(\mathbf{R}(t)) = \int_{t_i}^{t_f} L(\mathbf{R}(t), \dot{\mathbf{R}}(t), t)dt \tag{2.26}$$

The Principle of least action states that the path taken by the system between time t_i and t_f is the one in which the action is stationary to the first order - when $\partial S = 0$. The Euler-Lagrange equation can thus be derived as

$$\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{\mathbf{R}}_\mathbf{i}}\right) - \frac{\partial L}{\partial \mathbf{R_i}} = 0 \tag{2.27}$$

Since $\frac{\partial L}{\partial \dot{\mathbf{R}}_\mathbf{i}} = m_i \dot{\mathbf{R}}_\mathbf{i}$, and $\frac{\partial L}{\partial \mathbf{R_i}} = -\frac{\partial U}{\partial \mathbf{R_i}} = F_i$, the Euler-Lagrange equation is equivalent to Newton's equation of motion. The Lagrangian formulation holds in any coordinate system.

The Nosé thermostat

The Nosé approach [81] controls temperature using an extended Lagrangian which contains additional artificial coordinates and velocities. The fictitious dynamical variable, $\tilde{s}$ introduced in the Nosé scheme can be thought of as the position of an imaginary heat bath coupled to the system. $\tilde{s}$ has associated with it a conjugate momentum, $\dot{\tilde{p}}_s$ and heat bath mass, Q. Q determines the coupling between the heat bath and the system and influences the temperature fluctuation, whereas $\tilde{s}$ is a time-scaling parameter which stretches the timescale of the extended system, dt

$$d\tilde{t} = \tilde{s}dt \tag{2.28}$$

The Nosé method defines the Lagrangian for the extended system as:

$$L(\tilde{R}, \dot{\tilde{R}}) = \sum_i \frac{M_i}{2}\tilde{s}^2\dot{\tilde{\mathbf{R}}}_i^2 - U(\tilde{\mathbf{R}}) + \frac{1}{2}Q\dot{\tilde{s}}^2 - gk_bT_0 \ln \tilde{s} \tag{2.29}$$

where g is related to the number of degrees of freedom. The first two terms in Eq. 2.29 correspond to the Lagrangian to the system defined in Eq. 2.25. The third and fourth term correspond to the kinetic and potential energies of the added degrees of freedom of the heat bath, respectively. The particle momentum is calculated as

$$\mathbf{p}_i \equiv \frac{dL}{d\dot{\mathbf{R}}_i} = M_i\tilde{s}^2\dot{\tilde{\mathbf{R}}}_i \tag{2.30}$$

and the momentum conjugate to $\tilde{s}$ is

$$p_s \equiv \frac{dL}{d\dot{\tilde{s}}} = Q\dot{\tilde{s}} \tag{2.31}$$

The Hamiltonian of the extended system can be expressed as

$$H_{Nosé} = \frac{1}{2}\sum \frac{\mathbf{p}_i^2}{M_i\tilde{s}^2} + U(\tilde{\mathbf{R}}) + \frac{p_s^2}{2Q} + gk_bT_0 \ln \tilde{s} \tag{2.32}$$

The simulated system ($\mathbf{R}$, $\mathbf{p}$, $\mathbf{t}$) and extended system ($\tilde{\mathbf{R}}$, $\tilde{\mathbf{p}}$, $\tilde{\mathbf{t}}$) have the same atomic coordinates leading to

$$\tilde{\mathbf{R}} = \mathbf{R}, \;\; \dot{\tilde{\mathbf{R}}} = \frac{\dot{\mathbf{R}}}{\tilde{s}}, \;\; \tilde{s} = s \;\text{ and }\; \dot{\tilde{s}} = \frac{\dot{s}}{\tilde{s}} \tag{2.33}$$

The particle momentum in the extended system is thus scaled as

$$\mathbf{p}_i = M_i \dot{\mathbf{R}}_i \times \tilde{s} \tag{2.34}$$

The Nosé thermostat leads to the following equations of motion

$$\begin{aligned} \ddot{\tilde{\mathbf{R}}}_i &= \frac{\tilde{\mathbf{F}}_i}{M_i \tilde{s}^2} - \frac{2\dot{\tilde{s}}\dot{\tilde{\mathbf{R}}}_i}{\tilde{s}}, \\ \ddot{\tilde{s}} &= \frac{1}{Q\tilde{s}} \left(\sum_i M_i \tilde{s}^2 \dot{\tilde{\mathbf{R}}}_i^2 - g k_b T_0 \right). \end{aligned} \tag{2.35}$$

These equations of motion sample a microcanonical ensemble in the extended system but return a canonical ensemble of the original system. The thermodynamics of the canonical ensemble approximate the true dynamics of the system because the time evolution of the particle positions is deterministic and does not involve stochastic changes.

The Nosé-Hoover thermostat

The Nosé thermostat is impractical to use because $\tilde{s}$ is a variable throughout the simulation. Eq. 2.34 implies that the simulation time which should be used to evaluate kinetic properties and time-averages is also a variable. The Nosé-Hoover thermostat [82] is a reformulation by Hoover that overcomes these issues. In the Nosé-Hoover scheme p_s is replaced with a friction coefficient, ξ which slows down or accelerates particles until the desired temperature is reached. The Nosé-Hoover approach reformulates the equations of motion in terms of the real system variables using

$$\begin{aligned} s = \tilde{s}, \;\; \dot{s} = \tilde{s}\dot{\tilde{s}}, \;\; \ddot{s} = \tilde{s}^2\ddot{\tilde{s}} + \tilde{s}\dot{\tilde{s}}^2, \\ \mathbf{R} = \tilde{\mathbf{R}}, \;\; \dot{\mathbf{R}} = \tilde{s}\dot{\tilde{\mathbf{R}}}, \;\; \ddot{\mathbf{R}} = \tilde{s}^2\ddot{\tilde{\mathbf{R}}} + \tilde{s}\dot{\tilde{\mathbf{R}}}^2 \end{aligned} \tag{2.36}$$

The friction coefficient is defined as

$$\xi = \frac{\dot{s}}{s} \tag{2.37}$$

and the derivative is obtained using Eq. 2.36

$$\dot{\xi} = \frac{s\ddot{s} - \dot{s}^2}{s^2} = s\ddot{\tilde{s}} \tag{2.38}$$

which allows the Lagrangian equations of motion Eq. 2.35 to be rewritten as

$$\ddot{\mathbf{R}}_i = \frac{\mathbf{F}_i}{M_i} - \xi\mathbf{R}_i, \tag{2.39}$$

and

$$\begin{aligned} \dot{\xi} &= \frac{1}{Q}\left(\sum_i M_i\dot{\mathbf{R}}_i^2 - gk_bT_0\right) \\ &= \frac{gk_bT_0}{Q}\left(\frac{T(t)}{T_0} - 1\right) \end{aligned} \tag{2.40}$$

The Nosé-Hoover thermostat can be implemented by modifying the velocity Verlet algorithm in Eq. 2.12 and Eq. 2.13 . The thermostat relies on the choice of the heat bath mass Q. Too small values of Q can cause high-frequency temperature oscillations, whereas too large values result in poor temperature control and can only reproduce the canonical distribution after very long simulation times. In addition, in some cases the thermostat is not ergodic and can become trapped in subspace [83].

Barostats

A simulation can be performed at constant pressure in the NPT ensemble by using a barostat. The barostats are analogues of to their thermostat counterparts and treat the system volume as a dynamical variable. The simplest approach, volume rescaling, adjusts the system pressure to the target pressure by rescaling the system volume at periodic intervals.

The Berendsen barostat [78] is similar to the thermostat and couples the system to a pressure bath. An additional term is added to the equations of motion involving pressure

$$\frac{dP}{dt} = \frac{p_0 - p}{\tau_p} \tag{2.41}$$

where p_0 is the desired pressure, p the current pressure and τ_p a coupling constant. The Berendsen barostat scales the volume by η, and the particle coordinates and cell vectors by $\eta^{1/3}$

$$\eta = 1 - \frac{\beta \delta t}{\tau_p}(p_0 - p) \tag{2.42}$$

where β is the isothermal compressibility of the system. The Berendsen barostat suffers from similar limitations to the thermostat and cannot sample a true NPT ensemble.

The Nosé-Hoover barostat is analogous to the thermostat and drives the system to a state where the average internal pressure is equal to the applied bath pressure. The barostat includes an additional friction, η and modifies the equations of motion of the extended system to

$$\begin{aligned} \ddot{\mathbf{r}}_i &= \frac{\mathbf{F}_i}{M_i} - (\xi + \eta)\mathbf{R}_i, \\ \dot{\eta} &= \frac{V}{N k_b T_0 \tau_p^2}\left(p - p_0\right), \\ \dot{V} &= 3\eta V \end{aligned} \tag{2.43}$$

where τ_p is a specified time constant for pressure fluctuation, and p and V are the instantaneous system pressure and volume, respectively.

2.1.7 Limitations of Empirical Force Fields

Simulations using empirical force fields suffer from a number of limitations. No information relating to the electronic structure can be obtained and reactions such as bond breaking and forming, charge transfer and electron excitation cannot be simulated. The accuracy depends on the quality of the parameterisation and simulations are restricted to systems containing the functional group the force field was parameterised to. Simulations performed under conditions which differ significantly from the conditions of parameterisation can lead to errors.

The forces in force field molecular dynamics can be obtained relatively easily and quickly, in contrast to *ab initio* simulations which require solving the electronic structure to calculate forces. As a result, force field simulations are several orders of magnitude faster than *ab initio* simulations allowing much larger systems to be simulated for much longer time simulation times.

Another advantage of the force field based simulation approach is that it offers the possibility of modifying the potential to bias the simulation and reveal detailed information on how the properties are influenced by particular interactions of atomic contributions.

2.2 Enhanced Sampling Techniques

The efficiency of MD algorithms scales with the square of the number of particles. The time step used must be smaller than the fastest atomic motion to ensure numerical stability and accuracy of the simulations. The computational cost of MD simulations therefore restricts not only the size of systems that can be simulated but the time scales that are accessible. In general, studies using MD involve simulations of a few hundred nanoseconds but rarely extend into the microsecond regime and beyond. Nucleation and crystal growth do not occur within the nano time scale.

Such phenomena may occur in the millisecond or longer time scale - time scales that are inaccessible using MD simulations.

In an effort to improve the efficiency of MD simulations, coarse-grained models which describe a group of atoms as a single particle have been developed. This not only enables large systems to be modelled using fewer particles but can enable a larger time step to be used if the coarse-grained description removes the fast atomic bond vibrations by grouping bonded atoms into single particles.

Enhanced sampling techniques are modifications of MD which accelerate the dynamics of a simulated system, enabling much longer time scales to be accessed and accurate sampling of probability distributions, free energies and kinetic rate constants. Enhanced sampling techniques can help drive a system from one thermodynamic state to another by biasing the potential along a reaction coordinate, ξ. The reaction coordinate can be multi-dimensional and is a function of the atomic positions. The free energy along the reaction coordinate can be expressed using the probability distribution of the system along ξ

$$A(\xi) = -\frac{1}{\beta}\ln(Q(\xi)) \tag{2.44}$$

where β is the reciprocal of the thermodynamic temperature. A system in which every point in phase space is visited during a simulation is referred to as ergodic. In the limit of infinite sampling of an ergodic system the ensemble average $Q(\xi)$ is equal to the time average $P(\xi)$

$$P(\xi) = \lim_{t\to\infty}\frac{1}{t}\int_0^t \rho[\xi(t^{'})]dt^{'} \tag{2.45}$$

where t corresponds to time and ρ is a measure of the occurrence of state ξ. By monitoring the probability distribution of the system along the reaction coordinate, $P(\xi)$, it is possible to obtain A(ξ) directly from an MD simulation.

In MD simulations the configurational space around an energy minimum is usually well sampled, while regions with higher energies are sampled rarely. Configurations that are separated by an energy barrier that is significantly larger than k_BT remain

unsampled in the timescale of MD simulations. In such cases where these rare events do not occur, the system is no longer ergodic and $A(\xi)$ cannot be calculated from the MD simulations. Enhanced sampling methods have been developed to enable the sampling of rare events and allow the free energies to be recovered.

2.2.1 Metadynamics

Metadynamics is a powerful technique that is capable of reconstructing the free-energy surface as a function of selected degrees of freedom or reaction coordinates, referred to as collective variables (CVs). The CVs are functions of the atomic coordinates, $\mathbf{R}$ and can describe the $3N$ dimensions of the system in the reduced dimensionality of the CV or reaction coordinate space. In this technique, the Hamiltonian of the system is modified by adding an external history-dependent bias potential as a function of the CVs. The external potential discourages the system from revisiting configurations that have already been sampled by depositing repulsive Gaussians along the trajectory in the CV space.

A collective variable $\mathbf{S}(\mathbf{R})$ can be defined as a function of the microscopic coordinates $\mathbf{R}$. The bias potential at time t can be written as

$$U_G(\mathbf{S}(\mathbf{R}),t) = \sum_{t'=\tau,2\tau,\ldots<t} \omega \exp\left(-\frac{(\mathbf{S}(\mathbf{R},t)-\mathbf{S}(\mathbf{R},t'))^2}{2\sigma^2}\right), \tag{2.46}$$

where $\mathbf{S}(\mathbf{R},t)$ is the value of the CV at time t and position $\mathbf{R}$, $\mathbf{S}(\mathbf{R},t')$ is the value at previous times t', ω is the Gaussian height, σ is the Gaussian width and τ is the frequency at which the Gaussians are added. ω, σ and τ influence the accuracy and efficiency of the free energy reconstruction. Large Gaussians will result in a fast exploration of the free energy surface, however, the error in the calculation of the energy profile will be large. Conversely, using small Gaussians or depositing Gaussians infrequently will produce an accurate free energy surface but will take a long time.

After a sufficiently long time the bias potential can be used to estimate the underlying free energy

$$U_G(\mathbf{S}(\mathbf{R}), t \to \infty) = -F(\mathbf{S}(\mathbf{R})) + C \tag{2.47}$$

This relation has been validated empirically through extensive testing [84] and comparison with results from other enhanced sampling methods [85, 86].

Metadynamics offers a number of advantages compared to other enhanced sampling techniques. The technique inherently explores low free energy regions first and does not require *a priori* knowledge of the energy landscape. Systems tend to escape minima by passing through the lowest free energy saddle points, therefore metadynamics can be used to explore new reaction pathways. Multiple interacting metadynamics simulations can be performed, in what is known as multiple walkers metadynamics, each contributing to a combined overall history-dependent potential. However, there are two main drawbacks to metadynamics. Firstly, U_G does not converge to the free energy but oscillates around it. Secondly, identifying a set of appropriate CVs can be far from trivial.

Well-tempered metadynamics is able to converge the bias potential to the free energy by decreasing the height of the Gaussian with simulation time according to

$$\omega = \omega_0 \exp(\frac{-U_G(\mathbf{S}(\mathbf{R}), t)}{k_B \Delta T}), \tag{2.48}$$

where ω_0 is the initial Gaussian height, ΔT is an input parameter in units of Kelvin. The bias potential converges to

$$U_G(\mathbf{S}(\mathbf{R}), t \to \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{S}(\mathbf{R})) + C, \tag{2.49}$$

where T is the temperature of the system. At this point the probability distribution of the CVs becomes

$$P(S) \propto \exp\left(-\frac{F(\mathbf{S})}{k_B(T + \Delta T)}\right). \tag{2.50}$$

$\Delta T \to \infty$ corresponds to standard metadynamics, while $\Delta T = 0$ corresponds to standard molecular dynamics. In between these limits ΔT can be tuned to regulate the

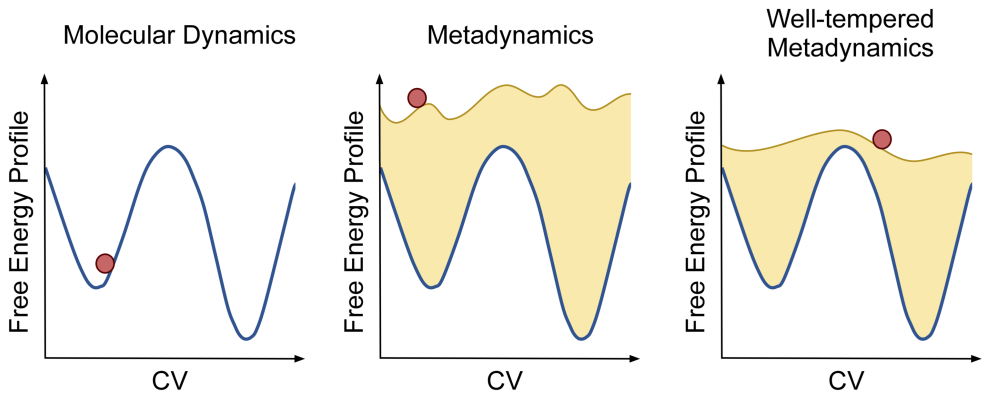


Figure 2.1: Schematic representation of a system's one-dimensional free energy landscape visited during MD, metadynamics and well-tempered metadynamics simulations. The true underlying free energy profile is plotted in blue whilst the red circle represents the current position of the system in the CV space. The bias potential is represented in yellow in the metadynamics plots. The well-tempered algorithm results in a more accurate convergence of the bias potential and prevents overfilling. Reproduced with permission from [87].

free energy surface exploration. Figure 2.1 is a schematic representation of molecular dynamics, metadynamics and well-tempered metadynamics in a one-dimensional free energy landscape. Well-tempered metadynamics enables a more accurate free energy to be reconstructed by preventing overfilling of the free energy basins with repulsive Gaussians.

Chapter 3

Simulating and Analysing Ice

This chapter introduces the models and parameters used for the molecular dynamics and metadynamics simulations in this thesis. Bespoke code has been developed to analyse the simulation trajectories and details of these codes are outlined in the following sections.

3.1 Force Fields for Ice

In order to perform molecular dynamic simulations of ice, it is necessary to choose a model of water. There is a vast number of water models to choose from because the overall properties of water cannot be fitted within a single set of parameters [88]. Specialised potentials have been developed and tailored for studies at particular temperatures and pressures, and for particular phases.

TIP4P and TIP5P are two rigid and non-polarisable models of water. TIP4P is a 4-site model while TIP5P is a 5-site model. Both models place a positive charge on the H atom and a Lennard-Jones interaction site on the O atom [89]. Where they differ is how they treat the negative charge on oxygen. TIP5P places a negative charge on two sites, representing the two lone-pair electrons of oxygen, whereas TIP4P models place the negative charge on a single site referred to as the M site. The M site is coplanar with the O and H atoms and is located at the bisector of the H–O–H angle [88], as shown in Figure 3.1. In TIP4P, the O–H distance and H–O–H

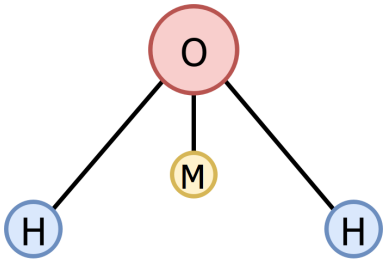


Figure 3.1: A schematic representation of the TIP4P 4-site model of water. There is one site for each atom and a 4^{th} M site models the negative charge of the oxygen atom.

angle are fixed to the experimental values of 0.9572 Å and 104.52 °, respectively [88]. A Lennard-Jones term and electrostatic interactions are the only contributions to the intermolecular pair potential of these models. Although the O atom bares no negative charge, it contributes to the Lennard-Jones term.

TIP4P/2005 is a re-parameterised and improved TIP4P model which was designed with the intention of producing a general model of water that is reliable for as many properties as possible. It is parameterised based on a fit to the temperature of maximum density and yields an excellent prediction of the density of water at 1 bar. It predicts a melting point at 1 bar of 252.1 K which is not in good agreement with the experimental value of 273.15 K, but is an improvement compared to 232 K predicted by TIP4P [88].

TIP4P/ICE is another adaptation of the TIP4P model which is designed to cope with the solid-phase interactions of water and to reproduce as closely as possible the phase diagram of ice. It has a predicted melting temperature for hexagonal ice at 1 bar of 272.2 K, which is in very close agreement with the experimental value. In addition, it has the best overall phase diagram and predicts most accurately the densities of several ice forms [90]. TIP5P predicts less successfully the phase diagram of water than the TIP4P/2005 and TIP4P/ICE models, despite being more computationally expensive [89].

In general, the more sites a model has, the more computationally expensive it is. Three site models are more computational efficient than the four-site TIP4P models. SPC/E [91] and TIP3P [92] are two examples of three site models which place a negative charge on the O atom and partial positive charges on the H atoms.

Intermolecular interactions between water molecules are computed using a Lennard-Jones type potential with a single interaction point centered on the oxygen atoms. Neither model includes van der Waals interactions involving hydrogen atoms and they differ only in their parameters. SPC/E is much poorer than the TIP4P models as it fails to successfully predict the phase diagram of water [88]. 3-site water models describe the radial distribution function poorly and TIP3P lacks the second peak in the Oxygen-Oxygen radial distribution function [93].

Modelling water atomistically requires at the very least a three-site model to take into account the three atoms in a water molecule. Water can equally be modelled using a single-site coarse-grained model which represents water as a particle with very short-ranged tetrahedral interactions. The mW model [94] is one example adapted from the Stillinger-Weber silicon potential. Water molecule interact through a Lennard-Jones Gaussian potential and a three body potential described by

$$
\begin{aligned}
E &= \sum_i \sum_{j>i} \phi_2(R_{ij}) + \sum \sum \sum \phi_3(R_{ij}, R_{ik}, \theta_{ijk}) \\
\phi_2(R_{ij}) &= A\epsilon \left[B \left(\frac{\sigma}{R_{ij}} \right)^4 - 1 \right] \exp\left(\frac{\gamma\sigma}{R_{ij} - a\sigma} \right) \\
\phi_3(R_{ij}, R_{ik}, \theta_{ijk}) &= \lambda\epsilon \left[\cos\theta_{ijk} - \cos\theta_0 \right]^2 \exp\left(\frac{\gamma\sigma}{R_{ij} - a\sigma} \right) \exp\left(\frac{\gamma\sigma}{R_{ik} - a\sigma} \right)
\end{aligned}
\tag{3.1}
$$

where R_{ij} is the distance between particles i and j, θ is the angle between three water molecules, and λ scales the repulsive three-body term and parameterises the strength of the tetrahedral interactions in the model. A, B and γ are constants, σ is the particle diameter, and $a = 1.8$ is a cut-off which ensures that all terms in the potential go to zero at a distance $a\sigma$. ϵ is the depth of the two-body interaction potential and is parameterised to the melting temperature, whilst the tetrahedrality λ is fitted to the vaporisation enthalpy. σ is fitted to the density of water at room temperature whilst the remaining parameters are taken from the original silicon model.

The mW model has no hydrogen atoms or electrostatic interactions and yet

successfully reproduces the energetics, density and structure of liquid water, in addition to its anomalies and phase transitions. The main advantage of the model, compared to atomistic models such as TIP4P, is its efficiency; mW is 2 orders of magnitude faster [94]. The model crystallises readily due to the absence of explicit hydrogen atoms, making it particularly suited to nucleation studies [95]. The simplicity of the model comes at a cost, however. It underestimates the density gap between ice and water and predicts a low isothermal compressibility. More significantly though, the model underestimates the heat capacity, C_p due to fewer degrees of freedom resulting in a degradation of the accuracy of the calculated energies and entropies [94].

In this project we use the TIP4P/Ice model of water as it has the best melting point prediction and reproduces the phase diagram of water more accurately than any other atomistic model. The authors who developed TIP4P/2005 and TIP4P/Ice models recommend the use of TIP4P/Ice in the studies of solid-water phases [88]. Simulations are also performed using the coarse-grained mW model to improve the efficiency of computationally expensive calculations and to provide comparisons and validations of TIP4P/Ice simulation results.

3.2 Ice Surface Models

The study of ice nucleation using molecular dynamics simulations is very challenging as the process is very slow and occurs on timescales that are inaccessible even with the most powerful supercomputers. In this project we overcome this issue by assuming a slab of ice has already formed and simulating the dynamics and energetics as it grows further.

Hexagonal ice can be described by an orthorhombic unit cell with mutually perpendicular axes [1]. For convenience this unit cell is used to build orthogonal ice slabs. The Buch algorithm [96] is used to disorder the proton configuration whilst enforcing the Bernal-Fowler rules [16]. This is achieved by initially building a perfect oxygen lattice. For each O–O axis, an H atom is randomly placed 1 Å

to one of the two oxygen atoms and the pair assigned as covalently bonded. At this stage each oxygen is covalently bonded to 0 to 4 hydrogen atoms. A Monte Carlo procedure is then performed to correct the initial H configuration so that eventually every oxygen atom is bonded to 2 hydrogens. The dipole moment of the system is then minimised by randomly selecting 10% of the O–O axes and randomly reassigning their hydrogens. The Buch algorithm is then repeated to disorder the proton configurations and ensure that each O atom is bonded to 2 H atoms. The resulting updated configuration is accepted if the dipole moment is reduced. A new random subset of O–O axes is then selected and the procedure is repeated until the dipole moment is less than 10^{-6} D. Two slabs are prepared for the simulations: one which exposes the basal face to vacuum and one which exposes the prism face to vacuum. The ice slabs consist of 5760 water molecules and are $\sim 93 \times 46 \times 44$ Å^3 in dimension. Figure 3.2 shows the basal and prism ice slabs.

Vacuum is inserted by elongating the dimension of the box by 40 Å in the direction perpendicular to the surface, producing two equivalent ice/vacuum interfaces, and resulting in a box of total size $\sim 133 \times 46 \times 44$ Å^3. Periodic boundary conditions are applied in all three directions and the size of the vacuum is chosen to be large enough (~ 40 Å) so that the two ice/vacuum interfaces do not interact. The length of the ice slab of over 90 Å ensures that the two surfaces are separated by sufficient bulk so they do not interact. Surface melting is anticipated during the simulations below the melting point but a slab of this length will ensure that the two premelted surfaces do not melt the entire slab.

The ice/vacuum interface mimics the ice/vapour interface at the low vapour pressure conditions we aim to study. 0.1 g/m^3 is close to the the upper limit of the vapour pressures we are interested in. Modelling such pressures in a volume of $40 \times 46 \times 44$ Å^3 requires a mass of $\sim 10^{-26}$ g. Given the mass of an oxygen atom is of the order of 10^{-23} g, modelling vapour at low pressures as vacuum is reasonable.

The ice/vacuum interface is approximately 16 molecular diameters in dimension in both directions along the interface, using a molecular diameter of water of ~ 2.75 Å [19]. Previous work estimates that about 10 molecular diameters are required for a

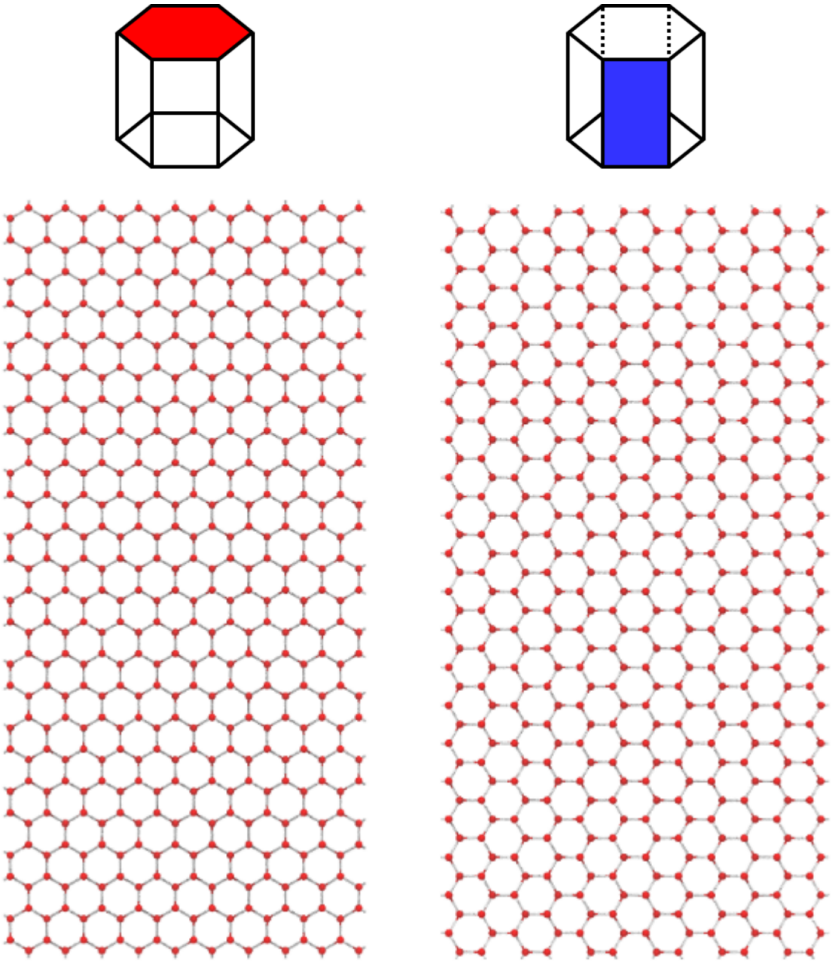


Figure 3.2: Initial configurations of the basal (left) and secondary prism (right) ice slabs used for TIP4P/ice simulations.

reliable estimate of the surface tension of the vapour-liquid interface [97]. Therefore, we should expect reliable dynamics and statistics from our simulations.

3.3 Molecular Dynamics Simulation Protocol

TIP4P/Ice

Molecular dynamic simulations are performed using the TIP4P/Ice potential at seven temperatures: 240 K, 245 K, 250 K, 255 K, 260 K, 265 K, and 270 K. It should be noted that the simulation temperatures are not absolute but relative to the melting point of the model. Since this is 272.2 K for TIP4P/ICE the seven temperatures that are simulated can be considered to be $\sim -2\,^{\circ}\mathrm{C}$, $-7\,^{\circ}\mathrm{C}$, $-12\,^{\circ}\mathrm{C}$, $-17\,^{\circ}\mathrm{C}$, $-22\,^{\circ}\mathrm{C}$,

−27 °C and −32 °C. The Molecular Dynamics package LAMMPS [98] is used for the simulations. The Lennard-Jones potential is truncated at 10.0 Å and the real part of the Coulombic potential is truncated at 12.0 Å. A particle-particle particle-mesh solver is used to compute long-range Coulombic and Lennard-Jones interactions in reciprocal space.

In order to equilibrate the ice slabs, 10-15 ns of Nose-Hoover NPT simulations are performed at zero pressure and at each temperature. A relaxation times of 100 fs and 1000 fs are used for the Nose-Hoover thermostat and barostat, respectively. The barostat allows the length of the two sides of the box that are in the plane of the exposed surface to be altered independently whilst keeping the box orthogonal. The third dimension along the length of the slab is not altered by the barostat because the presence of vacuum allows the system to expand or compress along the dimension of its own volition. A pressure of zero is used as an approximation to the very low atmospheric pressures within clouds that we aim to simulate. During these simulations the time step is gently increased from 0.01 to 1 fs for the first 3000 steps, followed by an incremental increase in the temperature for 4000 fs from 50 K up to the target temperature. 10 ns of NPT simulations ensure that the slab dimensions have equilibrated and the system relaxed.

Once the NPT simulations are complete the final NPT configuration is used to run NVT simulations. 10 ns of NVT simulations are run to equilibrate the system following by at least 90 ns of NVT simulations over which statistics are collected. A time step of 1 fs is used and the temperature is regulated using the Nose-Hoover thermostat with a relaxation time of 100 fs. During the NVT simulations, the dimensions of the simulation box are kept fixed.

mW

Molecular dynamics simulations of the basal and prism surfaces are performed using the mW potential at the same seven temperatures as TIP4P/Ice. Simulations are performed on slabs in which the hydrogen atoms have been removed leaving only the oxygen atoms. 10-15 ns of NPT simulations are performed using the Nose-Hoover

thermostat and barostat at zero pressure using a time step of 5 fs. Relaxation times of 1500 fs and 1000 fs are used for the barostat and thermostat, respectively. Once the NPT simulations are complete, 10 ns of NVT simulations are performed on the final NPT configurations to equilibrate the systems, followed by NVT runs of over 90 ns to collect statistics. A time step of 5 fs is used and the Nose-Hoover thermostat regulates the temperature with a relaxation time of 1000 fs. The five-fold increase in the time step, combined with the computational efficiency of the mW model due to the absence of long-range interactions, results in much more efficient simulations than with TIP4P/Ice, albeit at the cost of precision.

3.4 Analysing Simulations of Ice

In order to analyse the behaviour of the QLL that forms on the surface of ice and the surface dynamics of our simulations, it is necessary to define measures that can distinguish between ice-like molecules and liquid-like molecules. The third-order Steinhardt parameter and the Lindemann parameter are two different approaches that can achieve this which we have encoded for analysis of our simulations.

In order to elucidate the structural details of the QLL we have developed bespoke code to analyse the size of the ice-like and liquid-like clusters, the distributions of dangling bonds and the molecular orientation of molecules within our simulations.

3.4.1 Third-Order Steinhardt Parameter

Steinhardt bond-order parameters are sensitive to the degree of correlation of the spatial orientation of the vectors that join neighbouring particles [99]. Molecules are classified as crystal-like or liquid-like based on the coherence of their orientational order with that of their neighbours [100]. In ice, the tetrahedra centered on neighbouring molecules are aligned with a degree of order, while in water the alignment is random. This difference is the basis of the Steinhardt bond-order parameters used in this work.

The (local) Steinhardt order parameter of particle i, $q_l(i)$ is a normalised complex

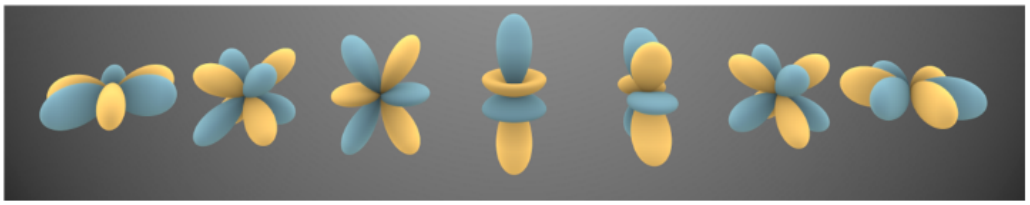


Figure 3.3: Schematic representation of the real Y_{3m} spherical harmonics, where m= -3 to +3. The function gives positive values in blue regions and negative values in yellow regions [103].

vector which identifies the orientation of particle i with respect to its neighbours. It has $(2l+1)$ complex components, $q_{lm}(i)$ which are defined as

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\mathbf{R}_{ij}) \tag{3.2}$$

where $N_b(i)$ is the number of nearest neighbours of particle i, l is an integer parameter, m is an integer that runs from $-l$ to $+l$, $\mathbf{r}_{ij}$ is the vector from particle i to its nearest neighbour particle j, and $Y_{lm}(\mathbf{r}_{ij})$ are the spherical harmonics [100, 101]. The $q_{lm}(i)$ terms project the orientational structure of the nearest neighbours of a molecule i on a basis of spherical harmonics. The Steinhardt order parameter, $q_l(i)$, can then be defined in terms of $q_{lm}(i)$ as [101]

$$q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2} \tag{3.3}$$

q_l parameters depend only on the angles between the vectors which connect neighbouring particles and are therefore independent of a reference frame [101]. They are rotationally invariant [102] and are sensitive to particular crystal symmetries depending on the choice of l. In this project, q_3 parameters are calculated as the $l = 3$ spherical harmonics can best describe tetrahedrality [102] and provide the best resolution for distinguishing between liquid-water molecules and hexagonal ice molecules [100]. Figure 3.3 shows the $l = 3$ spherical harmonics onto which the q_{3m} terms are projected. The $l = 3$ spherical harmonics are sensitive to tetrahedral symmetry and can be used to distinguish ice and water molecules because the Oxygen-Oxygen network has tetrahedral symmetry in ice but is disordered in water.

In order to improve the distinction between different structures using the q_l parameters, the following averaged form of the local Steinhardt order parameters are used in this project [101]

$$\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2} \tag{3.4}$$

where the q_{lm} vectors of particle i are replaced by the average over the nearest neighbours of particle i and particle i itself

$$\bar{q}_{lm}(i) = \frac{1}{N_b(i)+1} \sum_{j=0}^{N_b(i)} q_{lm}(j) \tag{3.5}$$

$\bar{q}_3(i)$ is calculated using the oxygen atoms only. While $q_l(i)$ contains information about the first shell around particle i, $\bar{q}_l(i)$ also takes into account the second shell [101]. This averaging enables a more accurate determination of the local crystalline structure due to the resulting narrower order parameter distribution [104].

In order to calculate $\bar{q}_l(i)$, the number of nearest neighbours of particle i, $N_b(i)$, must be known. $N_b(i)$ is determined using a neighbour cut-off distance. The neighbour cut-off distance used in the literature varies. 3.6 Å and 3.5 Å are used in [95] and [102], respectively, while [105] suggests a cut-off distance comparable to $1.2r_o$, where r_o is the first peak in the radial distribution function of the atoms concerned. In [106], 3.5 Å is used as it corresponds to the position of the first minimum of the oxygen–oxygen pair correlation function in the liquid phase. In this project we use 3.5 Å.

In order to determine a $\bar{q}_3$ threshold value which can be used to distinguish ice-like and water-like molecules, simulations of ice and water are performed using TIP4P/Ice at 260 K for 1 ns. q_3 and $\bar{q}_3$ are calculated from the simulations and the distributions in ice and water are plotted in Figure 3.4. The q_3 distributions of ice and water overlap significantly indicating that q_3 is a poor choice for identifying ice and water molecules. However, the overlap of the ice and water $\bar{q}_3$ distributions is small and therefore $\bar{q}_3$ can be used to accurately identify ice and QLL molecules in our simulations. Based on the point of intersection of the ice and water distributions,

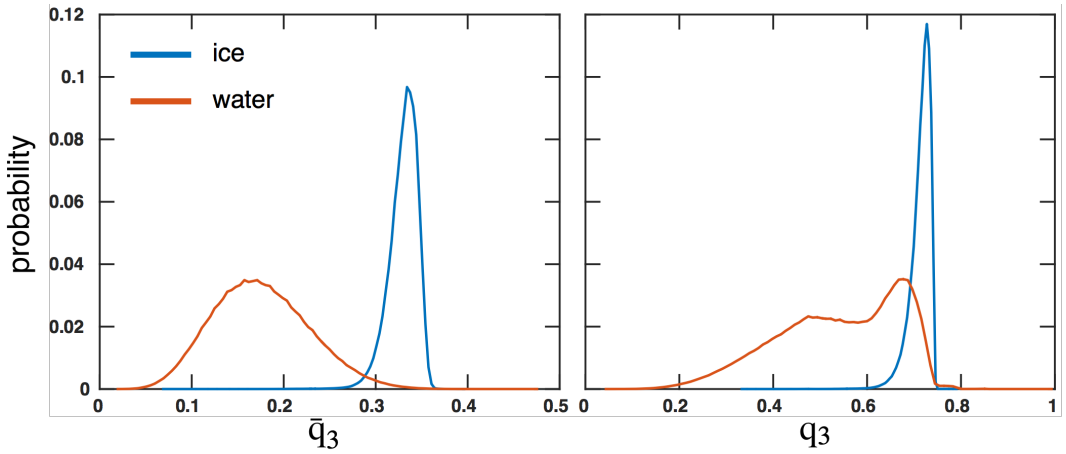


Figure 3.4: Distribution of $\bar{q}_3$ and q_3 calculated from MD simulations of ice and water at 260 K using TIP4P/Ice.

we defined a $\bar{q}_3$ threshold value of 0.28. An oxygen atom, and the H_2O molecule it corresponds to, is considered to be ice-like if $\bar{q}_3(i) \geq 0.28$ and liquid-like if $\bar{q}_3(i) < 0.28$.

From this point forward we refer to $\bar{q}_3$ simply as q_3 for simplicity, unless otherwise stated.

3.4.2 Lindemann Parameter

Analysis of the root mean square fluctuations in the O–O lengths can yield insight into the position of the QLL/ice interface and the QLL thickness [107]. This method is adapted from the Lindemann parameter.

The Lindemann parameter is based on Einstein's description of a crystal as vibrating harmonic oscillators. The parameter is defined as the ratio of the root mean square fluctuation in atomic positions about the equilibrium lattice positions to the nearest neighbour distance [108]. It is a measure of the root mean square vibration amplitude of a crystal and is defined for atom i as [109]

$$L_i = \frac{1}{N-1} \sum_{j \neq i} \frac{\sqrt{\langle R_{ij}^2 \rangle - \langle R_{ij} \rangle^2}}{\langle R_{ij} \rangle}, \tag{3.6}$$

where N is the number of water molecules or oxygen atoms, R_{ij} the distance between the oxygen atoms of molecule i and j, and $\langle \cdots \rangle$ refers to the time average.

As temperature increases, the amplitude of thermal vibration increases. Melting

occurs once the amplitude is large enough for adjacent atoms to partly occupy the same space. This is quantified by the Lindemann criterion which states that melting occurs when the dimensionless Lindemann parameter exceeds a critical value. This critical value varies between 0.05 and 0.2 depending on the crystal structure, magnitude of quantum effects and the nature of interparticle interactions [108]. For ice a value of 0.1 is indicated in [107] which suggests that melting of ice occurs once the amplitude of atomic thermal vibrations exceeds 10% of the atomic separation between neighbouring molecules. L_i can be calculated for every oxygen atom i and the profile plotted as a function of the coordinate normal to the surface. Any region satisfying $\langle L_i \rangle \geq 0.1$ can be regarded as belonging to the QLL and any O atom with $L > 0.1$ can be regarded as liquid-like.

3.4.3 Identifying Dangling Bonds

An OH bond is defined as dangling if it is not hydrogen bonded to any other water molecule. Dangling bonds tend to occur at a surface and result in polarisation [110]. They have a propensity to hydrogen bond and influence the dynamics at the surface through their favourable receptiveness to hydrogen-bond donor or receptor molecules. Dangling bonds are of interest in this work because they play an important part in vapour/QLL interactions. Analysis of the dangling bonds in our simulations reveals insights into the structure of the basal and prism surfaces and uncovers any dependence of hydrogen-bond surface structure on temperature. Luzar and Chandler identify H-bonds between two water molecules with an O—O separation of less than 3.5 Å and an OH–O angle less than 30° [111, 112]. We use this construct to identify dangling bonds in this work.

3.4.4 Describing Molecular Orientation

The orientation of molecules is studied using our code by calculating the Euler angles. The reference and molecular frames used to calculate the angles are shown in Figure 3.5. The reference frame is the x-, y- and z-axes of the simulation box, where z is normal to the surface and runs parallel to the length of the box. x and y

are the two in-plane directions and are orthogonal to one another and to the z-axis. The molecular frame is described by the X-, Y- and Z-axes in Figure 3.5, where the X-axis corresponds to the vector connecting the two hydrogen atoms. The Z-axis is the molecular dipole moment and corresponds to the vector connecting the midpoint of the X-axis and the oxygen atom. The Y-axis corresponds to the vector normal to the X- and Z-axes, whilst the line of nodes, N corresponds to the intersection of the xy and XY coordinate planes and is perpendicular to the zZ plane.

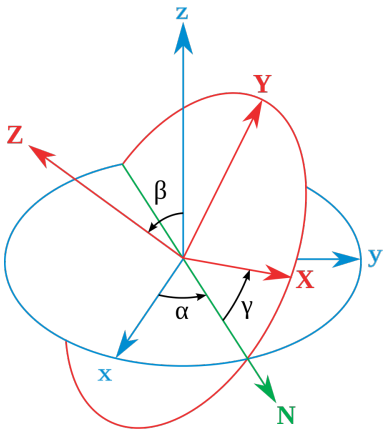


Figure 3.5: Schematic of Euler angles showing the reference xyz frame in blue and the rotated molecular XYZ frame in red. The line of nodes, N is shown in green (Image from Wikimedia Commons, Lionel Brits, 2008).

For our analysis, we calculate $\cos\beta$ corresponding to the angle between the dipole and the z-axis, and γ corresponding to the angle between the H-H vector and the line of node, N. Periodic boundary conditions are applied and N is calculated from the cross product of the dipole and the z-axis. γ captures the orientation of the hydrogen atoms in a molecule. When $\beta = 0$, a molecule is oriented with its dipole pointing directly to the surface.

3.4.5 Cluster Size

In order to study the QLL structure, we identify ice-like and liquid-like clusters in the QLL. In addition, we identify the largest ice-like cluster within our system which corresponds to bulk ice and use this cluster to determine the position of the QLL/ice interface.

A cluster is defined as a group of oxygen atoms which each lie within 3.5 Å

of at least one other oxygen atom from the cluster. Our algorithm relies on the assignment of molecules as ice-like and liquid-like based on the q_3 parameter and produces a global connection matrix which contains the connectivity of ice-like or liquid-like molecules. In order to identify a cluster, we use the depth-first search (DFS) algorithm. DFS is a recursive algorithm that is used to traverse a graph or tree. In our case the tree is the connectivity of a given atom and represents a cluster. The guiding principle of the algorithm is to progress forward in depth when possible and if not, backtrack and explore previous branches.

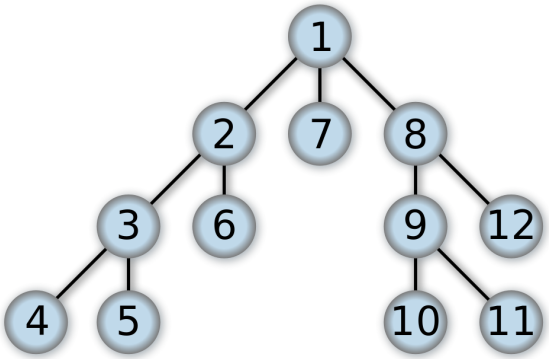


Figure 3.6: Schematic representation of the depth-first search (DFS) algorithm, illustrating cluster tree or oxygen connectivity. Each node represents an oxygen atom and the number indicates the order in which the atoms are visited in DFS (Image from Wikimedia Commons, Alexander Drichel, 2008).

Each oxygen atom corresponds to a node in the tree. Figure 3.6 illustrates a tree and shows the order in which DFS visits the nodes. A node can be in any of three states: *unvisited*, *visiting* and *visited*. Initially every node is *unvisited*. The algorithm works by arbitrarily choosing an *unvisited* node or vertex to analyse. At this point the cluster size count is one and the vertex is labelled as *visiting*. If the atom does not have any neighbours then the final cluster size count is one and the atom is labelled as *visited*. If, however, the atom v has at least one neighbour, the algorithm visits the first *unvisited* neighbour, u. The vertex atom v is labelled as *visiting* whilst the DFS algorithm is now performed recursively on the neighbouring molecule u which is also labelled as *visiting*. If u contains any *unvisited* neighbours, DFS is performed on the neighbours and thus the algorithm travels yet further down the tree. Once the algorithm encounters a molecule that has no *unvisited* neighbours,

the molecule in question is labelled as *visited*. The algorithm then backtracks to the previous molecule, p and visits any neighbours of p which are *unvisited*. Once all the neighbours of p are labelled *visited*, molecule p is also labelled *visited* and the algorithm backtracks yet again. This is repeated until every neighbour is *visited* and we have returned to the initial vertex v.

At this point, if the total atom count in the cluster is greater than five, our code saves the size of the cluster as well as the identity of the molecules in the cluster. If any *unvisited* atoms remain in the system, our algorithm randomly chooses a new *unvisited* vertex to perform DFS on, until every molecule is *visited*. Once this has completed we typically find ourselves with a number of saved clusters of size greater than five molecules which can be analysed further.

3.5 Metadynamics Simulation Protocol

3.5.1 Biasing Quasi-Liquid-Layer Molecules

Metadynamics simulations biasing the number of QLL molecules are performed using TIP4P/Ice and mW. Simulations are performed using LAMMPS [98] with the PLUMED [113] plugin. The CV used is the number of QLL molecule on the top surface, calculated using q_3. The CV is restricted to the top surface by limiting the range of z-coordinates over which it is calculated to a region of ~ 35 $\mathring{A}$ corresponding to the first 8 molecular bilayers. By restricting the CV to the top surface our simulations result in the melting and crystallisation of only one QLL surface enabling us to recover the free energy of the surface as a function of the number of QLL molecules. The q_3 parameter used in our analysis involves a sharp cut-off of 3.5 $\mathring{A}$ for identifying the nearest neighbours. In PLUMED, q_3 is calculated using a smooth cut-off described by a switching function of the form

$$s(R_{ij}) = \frac{1 - \left(\frac{R_{ij} - d_0}{R_0}\right)^{12}}{1 - \left(\frac{R_{ij} - d_0}{R_0}\right)^{24}} \tag{3.7}$$

where $R_0 = 3.5$ Å, $d_0 = 0$ and R_{ij} is the distance between atom i and j. $s(R_{ij})$

is parameterised to be set equal to one when atom j is in the first coordination sphere of atom i and is zero otherwise. $s(R_{ij})$ is used in PLUMED to ensure that the derivative of the CV is a continuous function.

NVT metadynamics simulations are performed using mW with a time step of 5.0 fs. The Nose-Hoover thermostat is used with a relaxation time of 1000 fs. Gaussians with a height of 2.0 kJ and $\sigma = 10$ are deposited every 2000 fs. The σ parameter is selected by monitoring the fluctuation of the CV during an unbiased run and setting it equal to between 1/3 and 1/2 of the fluctuations. A value of 10 is determined based on such considerations. A well-tempered bias factor of 50 is used and simulations of the basal and prism surfaces performed at 240 K, 260 K, 265 K and 270 K.

The CV used relies on a calculation of q_3 for the atoms and is computationally expensive and therefore metadynamics simulations of the basal and prism TIP4P/Ice surfaces are performed only at 260 K. NVT simulations are performed with a time step of 1.0 fs and using the Nose-Hoover thermostat with a relaxation time of 100 fs. Gaussians with a height of 2.0 kJ and $\sigma = 10$ are deposited every 400 fs. A well-tempered bias factor of 50 is used.

The starting configuration used for the mW and TIP4P/Ice metadynamics simulations are the final structures from the NVT MD simulations at the relevant temperatures. Metadynamics simulations are performed for up to 600 ns using mW and up to 350 ns using TIP4P/Ice.

3.5.2 Biasing Adsorption and Diffusion

Metadynamics simulations are performed in which the z-coordinate of a selected water molecule is biased. The CV used is the z-coordinate of a selected oxygen atom and the simulations enable the absorption and desorption of the molecule to be sampled along with its diffusion within the QLL and ice sublayers. Whilst the z-coordinate of the molecule is biased, the in-plane x- and y-coordinates are not. However, all three coordinates are restricted to limit the CV space that the molecule may explore and increase the rate of convergence. This is achieved using a restraining potential which starts acting on the system when the x-, y- or z-coordinates of the

biased molecule are greater or lower than defined threshold values, a. The restraining potential acts as a wall that restricts the coordinates of the molecule and has the general form

$$U_{res} = \kappa \left(s - a\right)^2 \tag{3.8}$$

where s is the coordinate of the molecule and $\kappa = 500$ kJ/mol is the force constant of the wall. The walls are positioned so that the in-plane coordinates of the biased molecule are restricted to exploring up to 15 Å in the x- and y- directions. An upper wall is positioned in the vacuum along the z-direction halfway between the two slab surfaces to prevent the molecule from diffusing to the second surface. A lower wall is positioned at the fourth bilayer within the bulk to prevent the molecule from exploring deep ice bilayers and force the molecule to explore the QLL and first few ice bilayers. The outcome of the walls is restricted in-plane diffusion and exploration of the CV space resulting in faster convergence of the reconstructed free energies.

NVT metadynamics simulations of the basal and prism surfaces are performed using TIP4P/Ice at 260 K. A time step of 1.0 fs is used along with the Nose-Hoover thermostat with a relaxation time of 100 fs. Gaussians with a height of 0.5 kJ and $\sigma = 0.075$ nm are deposited every 400 fs. σ is determined based on the fluctuations of the z-coordinate of unbiased QLL molecules. A well-tempered bias factor of 15 is used. The simulations are performed using the final structure from the NVT MD TIP4P/Ice simulations at 260 K. The metadynamics simulations are performed for 1100 ns.

Chapter 4

Simulating the Quasi-Liquid-Layer using Molecular Dynamics

Molecular dynamics simulations of the basal and prism surfaces of ice are performed using TIP4P/Ice and mW. The simulation protocols are described in Chapter 3.3. Snapshots of the TIP4P/Ice basal and prism ice slabs following 100 ns of simulation are shown in Figure 4.1. The formation of a disordered QLL is clearly visible on the surfaces. As temperature increases, the QLL thickness increases. At low temperatures there is one single melted layer, an observation consistent with recent experimental and simulation results [66]. At higher temperatures a second melted layer occurs, however, the temperature at which it is onset is unclear. Sanchez

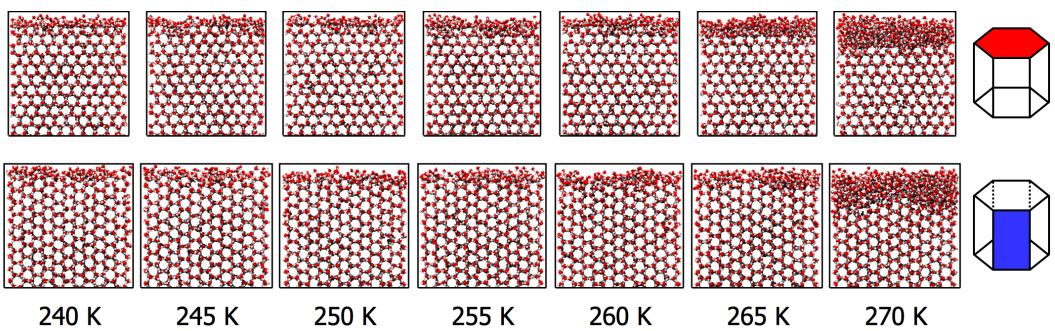


Figure 4.1: Snapshots of the TIP4P/Ice basal (top) and secondary prism (bottom) surfaces during NVT molecular dynamics simulations.

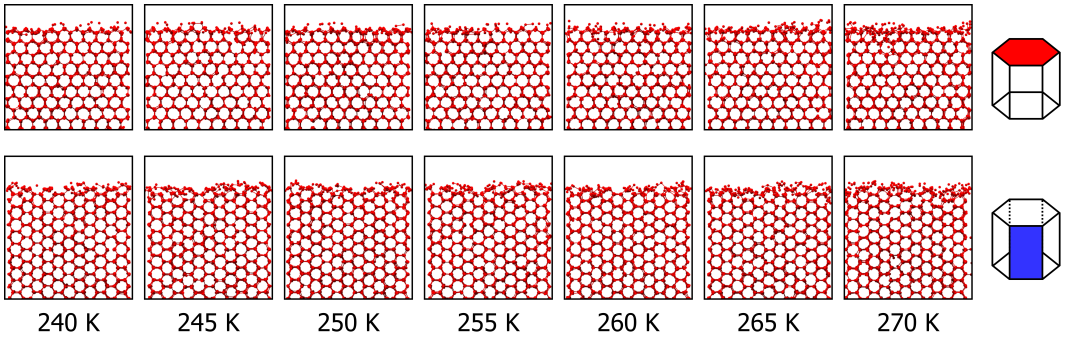


Figure 4.2: Snapshots of the mW basal (top) and secondary prism (bottom) surfaces during NVT molecular dynamics simulations.

[66] suggests a transition from one melted bi-layer to two melted bi-layers at 257 K based on SFG spectroscopy results.

Snapshots from the MD simulations of the basal and prism surfaces using mW are shown in Figure 4.1. The mW surfaces are over-structured compared to TIP4P/Ice. The mW QLL is composed of less than one bilayer and does not diverge in thickness as the melting point is approached, in contrast with our TIP4P/Ice simulation results and previous experimental [33–35, 66] and simulation works [7, 36–38].

The oxygen atom density profiles of the TIP4P/Ice basal and prism slabs are shown in Figures 4.3 and 4.4, respectively. The profiles are plotted along the z-axis

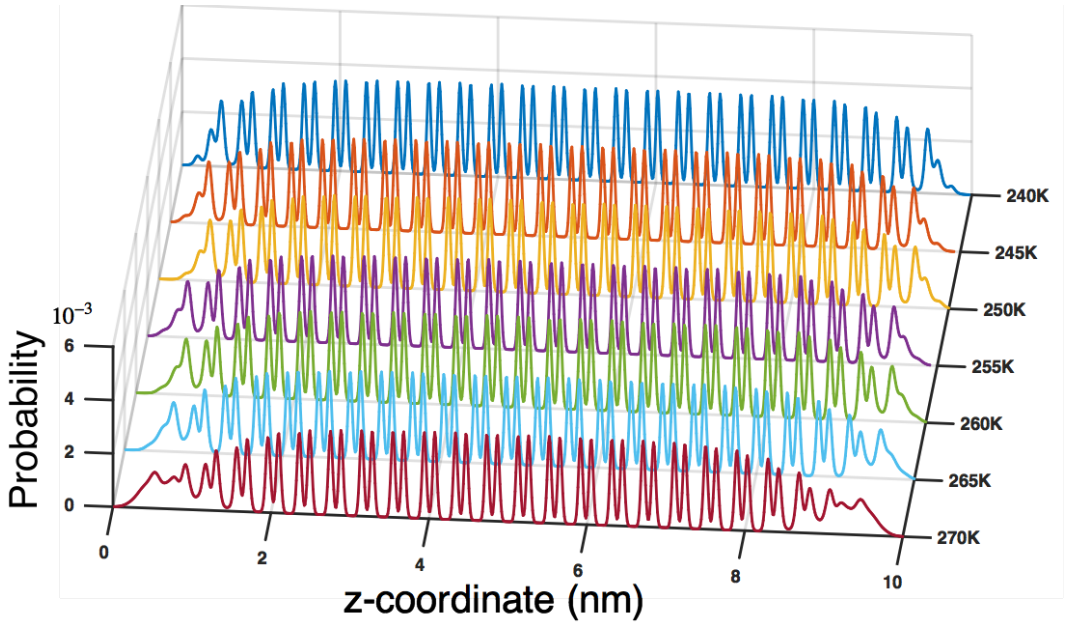


Figure 4.3: Oxygen density profiles of the basal TIP4P/Ice slabs at the seven simulated temperatures.

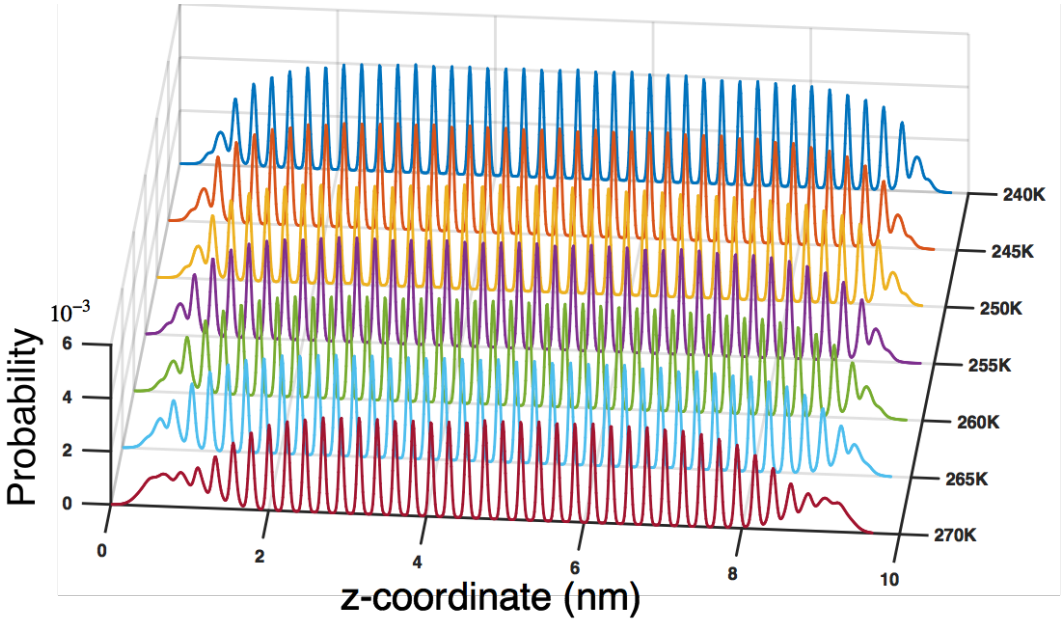


Figure 4.4: Oxygen density profiles of the (secondary) prism TIP4P/Ice slabs at the seven simulated temperatures.

normal to the surface and the coordinates are averaged over the entire simulation following 10 ns of equilibration. The profiles are calculating using 1000 bins of equal width. The plots reveal that as the temperature increases from 240 K to 270 K, the density peaks near the surfaces become broader and merge. This can be rationalised as the formation of a QLL with a thickness increasing with temperature. The plots show that there is always a central region where the peaks are sharp and the oxygen atoms localised, corresponding to bulk-crystalline ice.

The density profiles illustrate the general trend of increasing QLL thickness with temperature. However, more complex analysis is required to characterise the structure and thickness of the basal and prism QLLs.

4.1 Quasi-Liquid-Layer Size

The third-order Steinhardt bond-order parameter, $q_3(i)$ is calculated for every atom i using Eq. 3.4. Figures 4.5 and 4.6 show the q_3 profiles plotted against the coordinate perpendicular to the surface for the TIP4P/Ice and mW surfaces,

respectively. The profiles are averaged over the simulations following 10 ns of equilibration.

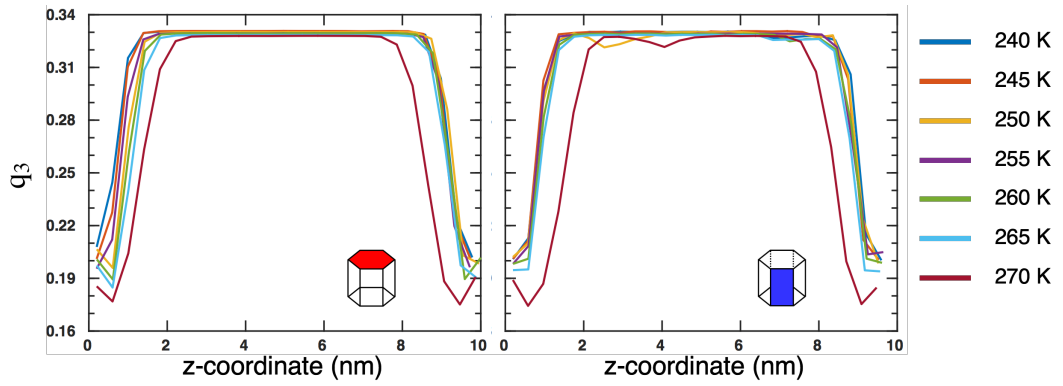


Figure 4.5: q_3 profiles of the basal (left) and prism (right) TIP4P/Ice surfaces. Large q_3 values corresponds to ice-like molecules and small values correspond to liquid-like molecules.

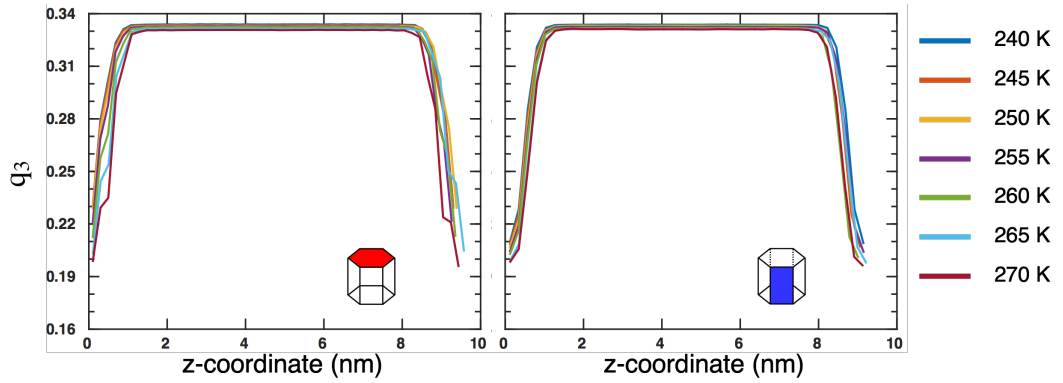


Figure 4.6: q_3 profiles of the basal (left) and prism (right) mW surfaces.

The profiles show that the q_3 parameter can be used to differentiate between ice-like and liquid-like molecules. In the bulk the q_3 value is ~ 0.32, whereas in the surface it is smaller at ~ 0.20. In this work we use a threshold value of 0.280 to distinguish between liquid-like and ice-like molecules based on the overlap of the q_3 distribution in water and ice.

In order to study the effect of temperature on the QLL thickness, we use q_3 to estimate the number of QLL molecules, n_l by classifying molecules as ice-like and liquid-like. The time evolution of the TIP4P/Ice basal and prism QLL size is captured by a plot of n_l against simulation time in Figure 4.7. The results show that the QLL fluctuates in size around an average value but as temperature increases, the average value and amplitude of fluctuations increase. This behaviour is consistent

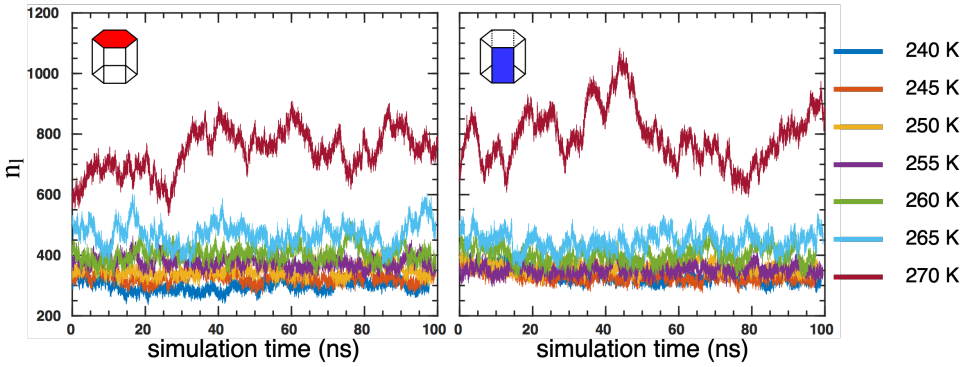


Figure 4.7: Evolution of the number of QLL molecules, n_l during the basal (left) and prism (right) TIP4P/Ice MD simulations.

with MD results reported with TIP4P/Ice and other water models [7, 48]. At 270 K, close to the TIP4P/ice melting point of 272.2 K [90], the QLL size fluctuates strongly as the outer ice-sublayer and inner QLL-sublayer readily melt and crystallise at the expense of one another. Figure 4.7 suggests that 100 ns simulations are too short to accurately sample the basal and prism QLLs at 270 K. As a result, the MD simulations at 270 K are extended to over 400 ns, and Figure 4.8 shows n_l for the basal and prism surfaces over the entire simulations. The results show the QLL fluctuating around equilibrium with large amplitudes.

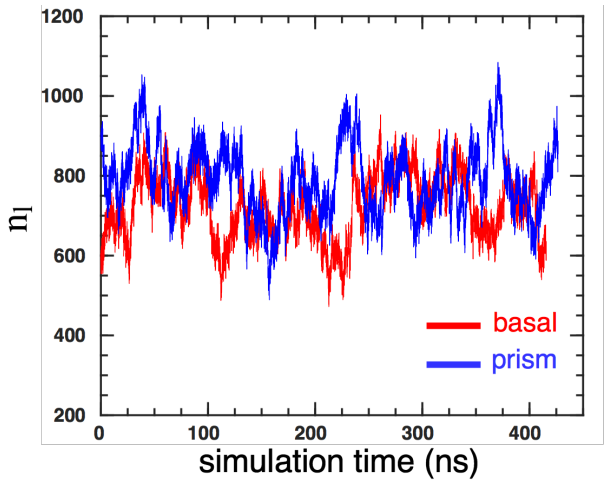


Figure 4.8: Evolution of the number of QLL molecules, n_l during the basal and prism simulations at 270 K using TIP4P/Ice

The evolution of n_l observed with mW is plotted in Figure 4.9 for the basal and prism surfaces. The mW QLLs fluctuate around an equilibrium thickness. Compared to TIP4P/Ice, the mW n_l fluctuations are smaller and do not increase

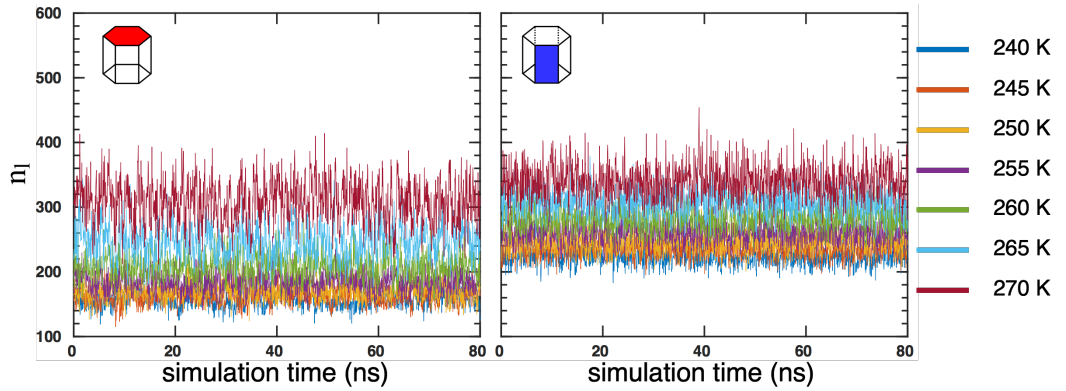


Figure 4.9: Evolution of the number of QLL molecules, n_l during the basal and prism simulations using mW

with temperature to the same extent.

The average basal and prism QLL size, $\bar{n}_l$ is estimated from our simulations and plotted as a function of temperature for the TIP4P/Ice and mW surfaces in Figure 4.10. The error bars indicate the standard deviation of n_l during the simulations and capture the trend of increasing amplitude of QLL size fluctuation with temperature observed in Figure 4.7. The results show that the QLL size increases with temperature irrespective of the water model. The TIP4P/Ice trend agrees with previous experimental results and can be described by the functional form $n_l \propto \log(\Delta T)$ [1, 33–35].

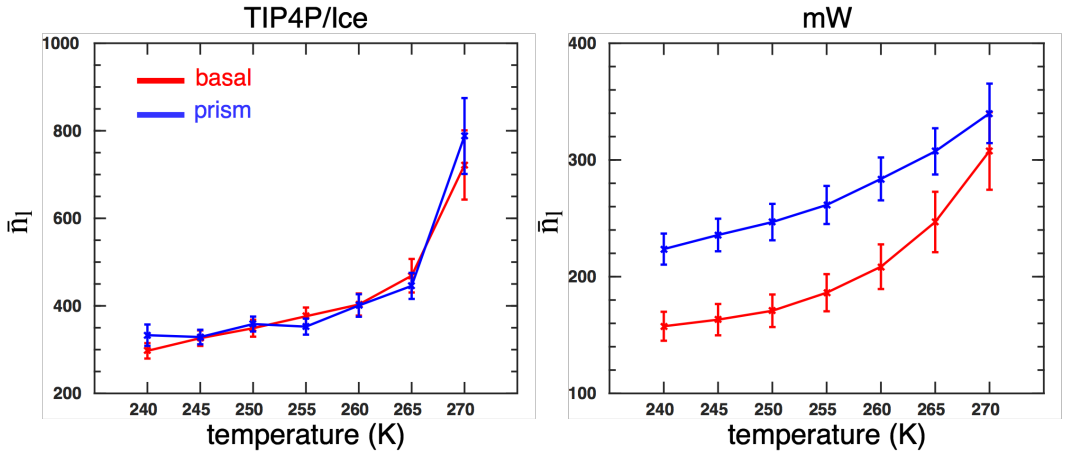


Figure 4.10: Average QLL thickness, $\bar{n}_l$ from MD simulations using (left) TIP4P/Ice and (right) mW.

The logarithmic behaviour of the TIP4P/Ice QLL size is investigated by fitting the basal and prism $\bar{n}_l$ results to $\bar{n}_l = C_1 + C_2 \log(273.15 - T)$, where C_1 and C_2

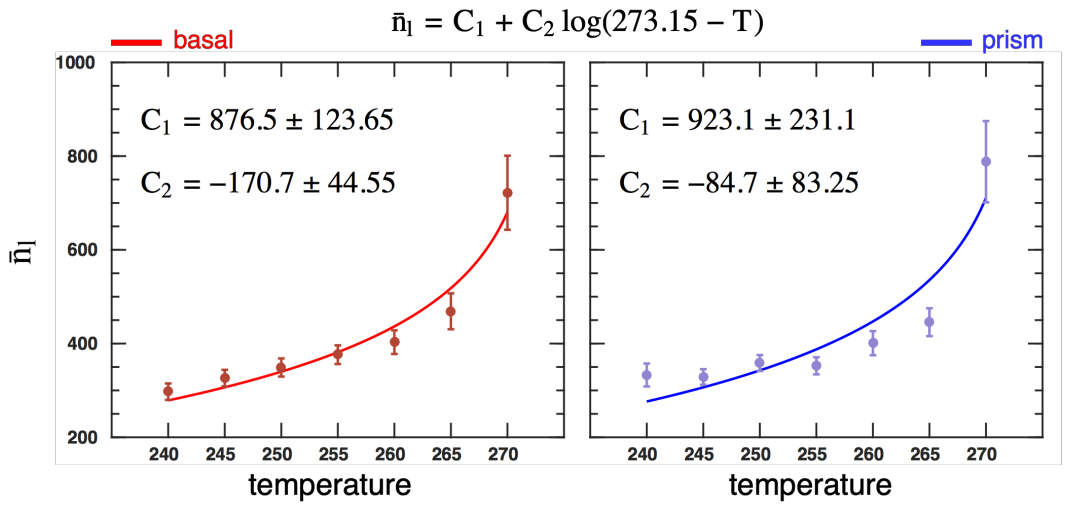


Figure 4.11: Logarithmic fit of $\bar{n}_l$ to temperature for the basal (left) and prism (right) TIP4P/Ice surfaces.

are fitting parameters. The results of the fits are shown in Figure 4.11. The fitting yields $C_1 = 876.5 \pm 123.65$ and $C_2 = -170.7 \pm 44.55$ for the basal surface, and $C_1 = 923.1 \pm 231.1$ and $C_2 = -184.7 \pm 83.25$ for the prism surface. The errors represents the 95% confidence interval. The basal and prism C_1 and C_2 fitting parameters overlap within the confidence bounds. The basal and prism TIP4P/Ice results suggest that the QLL thickness is dependent on temperature but depends weakly of the crystallographic surface.

The average q_3 and $\cos\beta$ values of the QLL molecules are calculated during our simulations and plotted as a function of temperature in Figure 4.12. Error bars indicate the standard deviation. Within the upper boundary of the standard deviation, the average q_3 of the QLL molecules is less than the threshold value of 0.28 which validates our choice of threshold value. β is the angle of the molecular dipole with the z-axis and is described in Section 3.4.4. The average $\cos\beta$ result in Figure 4.12 suggests that the local environment of QLL molecules changes very little between 240 K and 265 K and that the two crystallographic QLL surface structures are similar. This observation is consistent with previous vibrational response experiments which suggest that the outermost surface structure of ice remains unchanged between 235 K and 264 K [66, 69].

In contrast to TIP4P/Ice, mW is over-structured and a QLL of less than one

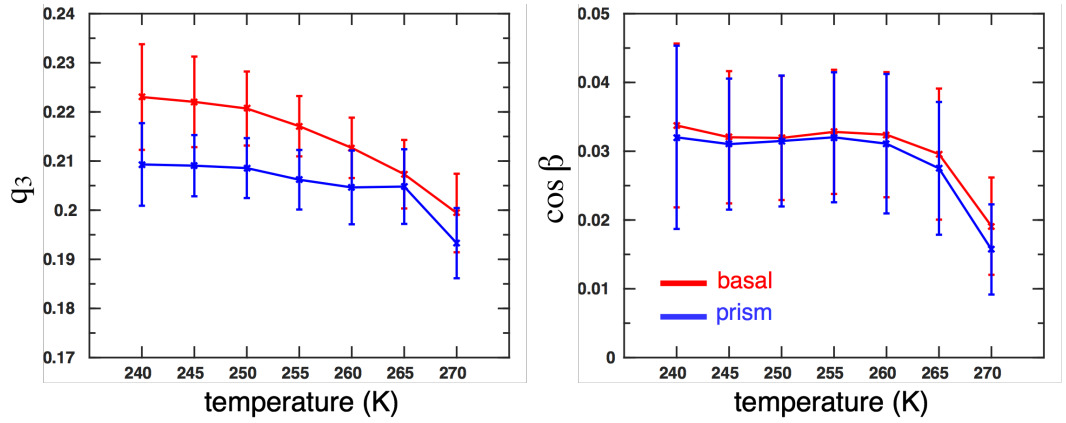


Figure 4.12: Average a) q_3 and b) $\cos\beta$ of the TIP4P/Ice QLL molecules

molecular layer is formed, as shown in the snapshots in Figure 4.2. 270 K is close to the melting point of the mW model reported at 274.6 K [94]. Whilst the mW QLL increases in size with temperature, it does not diverge at 270 K as strongly as the TIP4P/Ice QLL. The TIP4P/Ice QLL size exhibits a very weak dependence on crystallographic plane, whereas the mW prism QLL is larger than the basal QLL at all the simulated temperatures. Overall, the mW model gives a poor description of the QLL and does not fully characterise the QLL temperature dependence. In this work, we focus our studies of the QLL structure on results from TIP4P/Ice simulations. This decision is validated by the agreement of our TIP4P/Ice simulation results with the experimental QLL temperature dependence and surface structure, and agreement with previous simulation works [1, 7, 29, 33–38].

4.2 Lindemann Theory of Melting

The temperature dependence of the QLL can be explained using the Lindemann theory of melting. The Lindemann criterion states that melting occurs when the amplitude of thermal vibrations is greater than a threshold value. By modelling molecular vibrations using a harmonic potential, the amplitude of intermolecular fluctuations can be approximated as

$$x = \left(\frac{k_B T}{\kappa}\right)^{1/2} \tag{4.1}$$

where k_B is the Boltzmann constant, κ is an effective spring constant. Molecular bonding is weaker at the surface than in the bulk which is captured in

$$\kappa(d) = \kappa_0 - f(d)\Delta\kappa \tag{4.2}$$

where d is the distance from the surface, $\Delta\kappa$ is a constant, and $f(d)$ is a switching function with $f(0) = 1$ and $f(\infty) = 0$. The Lindemann criterion states that melting occurs once the amplitude x reaches $x_m = (k_B T_m/\kappa_0)^{1/2}$. Combining this with Eq. 4.1 and 4.2 yields

$$x(d) = \left(\frac{kT}{\kappa_0 - f(d)\Delta\kappa}\right)^{1/2} = x_m \tag{4.3}$$

Modelling the intermolecular interactions using $f(d) = \exp(-d/\lambda)$, where λ is a decay length which quantifies the fact that interface molecules are more weakly bound than bulk molecules, yields a QLL thickness described by

$$d = \lambda \ln\left(-\frac{\Delta\kappa T_m}{\kappa_0 \Delta T}\right) \tag{4.4}$$

where $\Delta T = T_m - T$ is the temperature below the melting temperature T_m. The functional form of Eq. 4.4 correlates with our TIP4P/Ice simulation results in Figure 4.11.

Lindemann Profile

The Lindemann parameter, outlined in Section 3.4.2, characterises the crystallinity of molecules in our simulations and is used to distinguish the QLL and ice sublayers. The Lindemann parameter is calculated for every O atom i, using Eq. 3.6, and the TIP4P/Ice basal and prism profiles are plotted as a function of the coordinates perpendicular surface in Figure 4.13. The parameters are calculated by averaging over the last 90 ns of the NVT simulations. The Lindemann parameter increases linearly with simulation time. As a result, the exact same number of frames are used for the analysis at the different temperatures to enable fair comparisons. The results show that in the bulk, the Lindemann parameters are small due to the localised

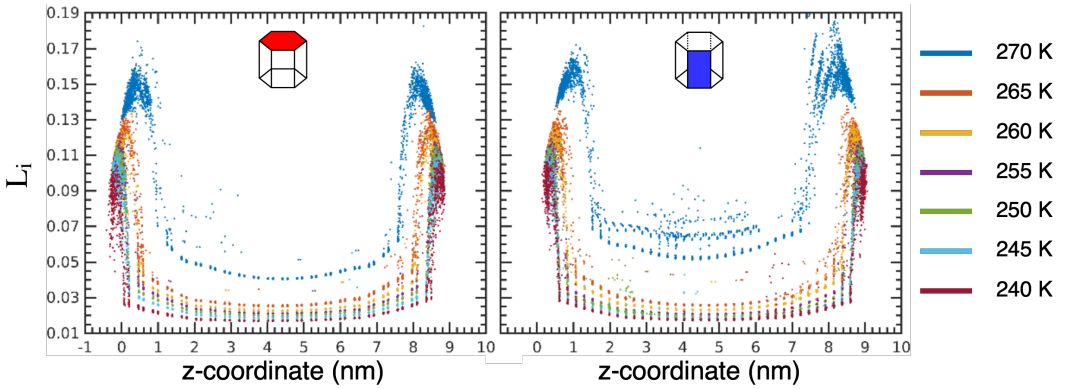


Figure 4.13: Lindemann profiles of the basal (left) and prism (right) TIP4P/Ice surfaces.

atomic vibrations with restricted amplitudes. The prism profile at 270 K suggests the presence of defects and ice molecules with increased mobility within the bulk. At the surface, the Lindemann parameters are much larger indicating the presence of the QLL where atomic motion is increased. From the profiles it is clear that the QLL region increases in size with temperature. As temperature increases the Lindemann parameter increases both in the bulk and at the surface. This is captured in Figure 4.14 which shows the minimum and mean Lindemann parameter calculated from the profiles and plotted against temperature. The general trend is similar to the TIP4P/Ice QLL thickness result shown in Figure 4.10: the basal and prism Lindemann parameters are equivalent and increase with temperature before diverging at 270 K.

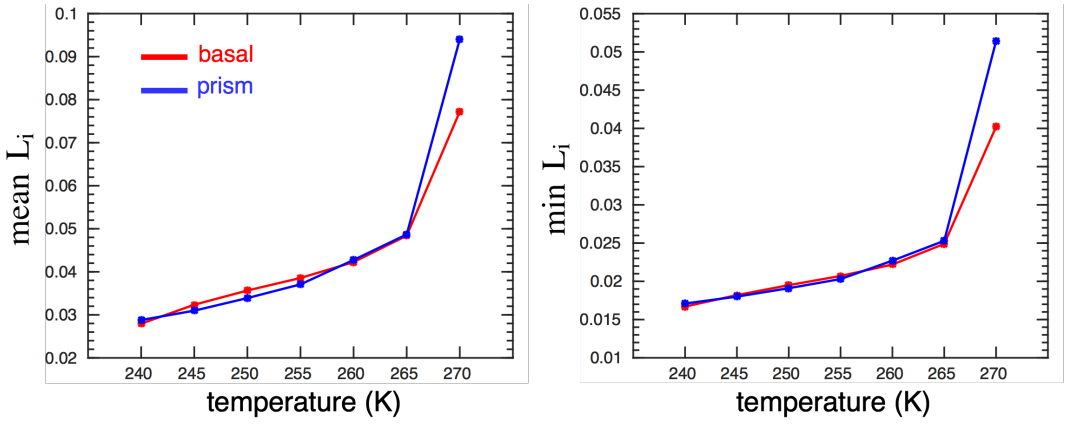


Figure 4.14: Temperature dependence of the mean (left) and minimum (right) Lindemann parameter for the TIP4P/Ice slabs.

The Lindemann profiles separate the ice-like and liquid-like regions of our slabs.

However, the parameter of ice and liquid molecules increase with temperature. As a result, the Lindemann parameter is less useful than q_3 for characterising the QLL size since the q_3 parameter of ice-like and liquid-like molecules does not depend on temperature, as can be observed in q_3 profiles in Figures 4.5 and 4.6

The density, q_3 and Lindemann profiles reveal a consistent picture. The surface of ice is disordered and the bulk ice is ordered. The thickness of the disordered QLL region depends on the temperature. However, the analysis does not reveal any significant differences between the basal and prism QLLs.

4.3 Characterising the Quasi-Liquid-Layer

4.3.1 Ice and Water Clusters

Insight into the QLL structure is obtained from the distribution of the sizes of the largest liquid-like clusters during the simulations. Figure 4.15 shows the distributions obtained from our basal and prism TIP4P/Ice simulations. The clustering algorithm used is outlined in Section 3.4.5. Between 240 K and 265 K, the distributions are unimodal and increase in width with temperature. At 270 K the width of the distribution is significantly larger than at the other temperatures. The basal distribution is multimodal at 270 K, with two distinct peaks separated by ~ 250. Given that one basal crystal layer in our system is composed of 240 molecules, this peak separation suggests that at equilibrium the QLL fluctuates between melting

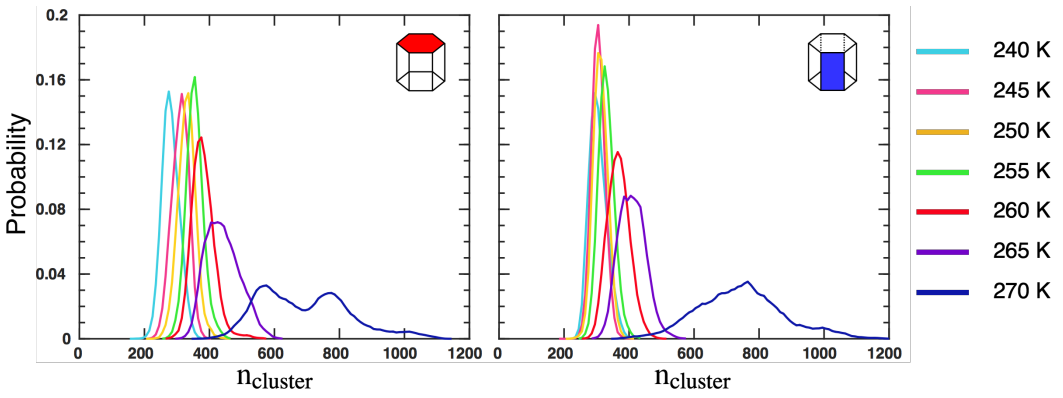


Figure 4.15: Largest liquid-like cluster size distribution for the basal (left) and prism (right) TIP4P/Ice surfaces.

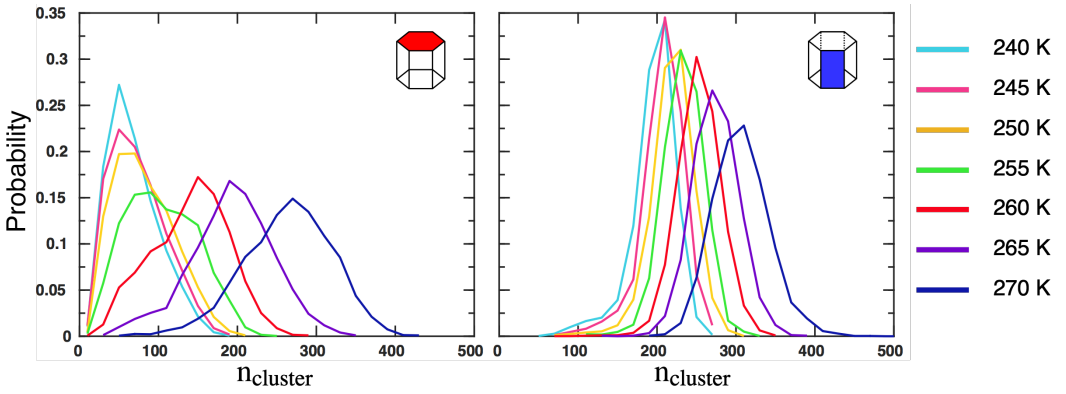


Figure 4.16: Largest liquid-like cluster size distribution for the basal (left) and prism (right) mW surfaces.

and crystallising two layers. The mW distribution of largest liquid-like clusters is plotted in Figure 4.16. The distributions are unimodal at all the temperatures.

The temperature dependence of the largest TIP4P/Ice and mW liquid-like cluster size is plotted in Figure 4.17. Error bars indicate the standard deviation over the simulations. The TIP4P/Ice trend is similar to the QLL thickness and mean and minimum Lindemann results: the cluster size increases with temperature and diverges at 270 K, and the basal and prism cluster sizes are equivalent within the standard deviation. This result is unsurprising as the largest liquid-like cluster corresponds to the cluster of liquid-like molecules in the QLL. The mW result strongly correlates with the mW QLL thickness result in Figure 4.10 and helps validate our observation

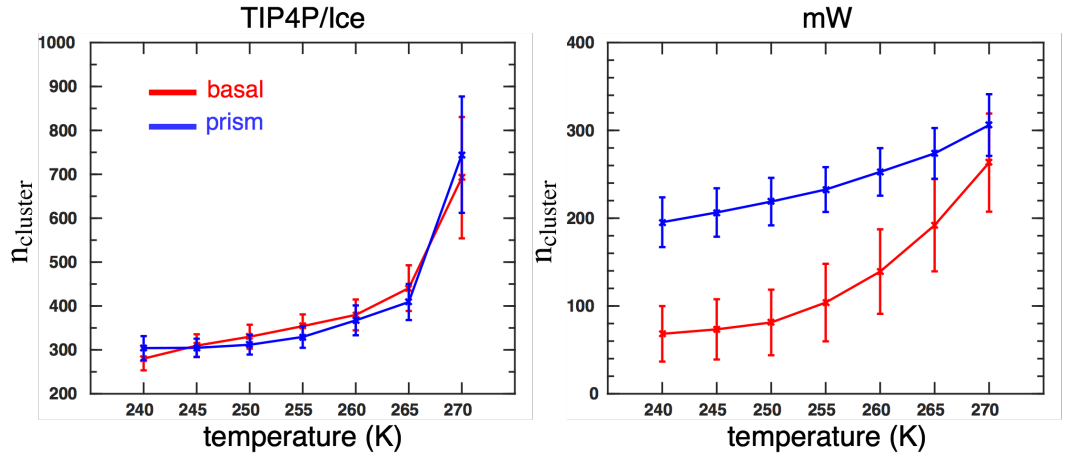


Figure 4.17: Largest liquid-like cluster size as a function of temperature for TIP4P/Ice (left) and mW (right) simulations.

that the QLL size does not diverge at 270 K to the same extent as with TIP4P/Ice. In principle, the largest liquid-like cluster size should correspond to the average QLL size. Contrasting the cluster sizes in Figure 4.17 with the QLL size in Figure 4.10, it is clear that the two are in close agreement. The small differences are due to the uncertainty associated with q_3 and QLL molecules incorrectly assigned as ice and ice molecules incorrectly assigned as QLL molecules.

4.3.2 Dangling OH bonds in the Quasi-Liquid-Layer

A QLL forms on the surface of ice to increase the stability. Dangling bonds may also form as the system endeavours to lower its surface energy. Analysis of the number of dangling molecules therefore reveals further insights into the structure of the QLL of ice. Figure 4.18 shows the distribution of the number of dangling molecules in our TIP4P/Ice basal and prism simulations. The peaks are unimodal and increase in both width and mean value as temperature increases. The basal and prism distributions appear very similar.

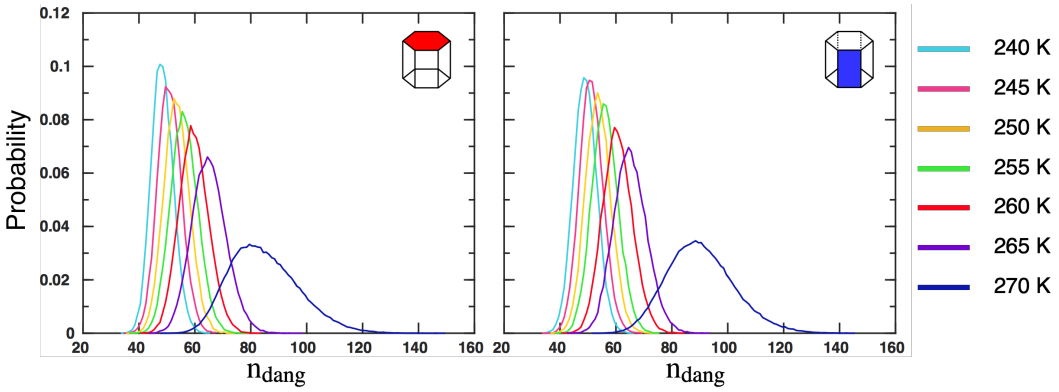


Figure 4.18: Distributions of dangling OH bonds on the a) basal and b) prism TIP4P/Ice surface

The number of dangling bonds per molecule is calculated from the number of dangling bonds divided by the number of QLL molecules. The average is calculated from the number of dangling bonds per molecule at each timestep and the result is plotted against temperature in Figure 4.19 with error bars indicating the standard deviation. The trend obtained is similar for the basal and prism surfaces. Between 240 K and 265 K there are ~ 0.15 dangling bonds per molecule which decreases to $\sim$

0.12 dangling bonds per molecule at 270 K. Assuming that the number of molecules with two dangling bonds is negligible, this result indicates that on average $\sim$ 12-15% of molecules are single hydrogen bond donors and have one dangling bond. The estimate for the QLL is averaged over the entire QLL and includes both surface QLL molecules and molecules belonging to the interior of the QLL. Within the interior of the QLL, the proportion of dangling molecules is expected to be lower than $\sim$ 12-15% and at the surface the proportion of dangling molecules is expected to be greater than $\sim$ 12-15%. Previous MD studies of water/air interfaces at 300 K show that 13% of molecules in the interior are single-donors which increases to 29% at the surface [114]. *Ab initio* studies affirm that 15% of bulk water molecules are single-donors whereas a 45% of surface molecules are single donors [115]. Our results are in good agreement with the bulk water estimates suggesting the QLL is mostly liquid-like as opposed to ice-like. However the interior QLL estimate is expected to be lower than the average QLL value. Perfect crystalline ice has no single-donor molecules [115] and therefore the number of single donors in the interior QLL lies between the values in bulk ice and bulk liquid implying that the QLL structure is influenced by the underlying ice phase. To help understand the environment of the basal and prism dangling bonds, Figure 4.19 also shows the average q_3 value and $\cos\beta$ associated with the dangling bonds. The results indicate that the orientation and environment of the dangling bonds depend on temperature but are independent of the crystallographic surface. The average q_3 suggests that between 240 K and 265 K, the bonding environment of dangling molecules is fairly consistent and is only different at 270 K when the QLL is substantially larger. This result is consistent with

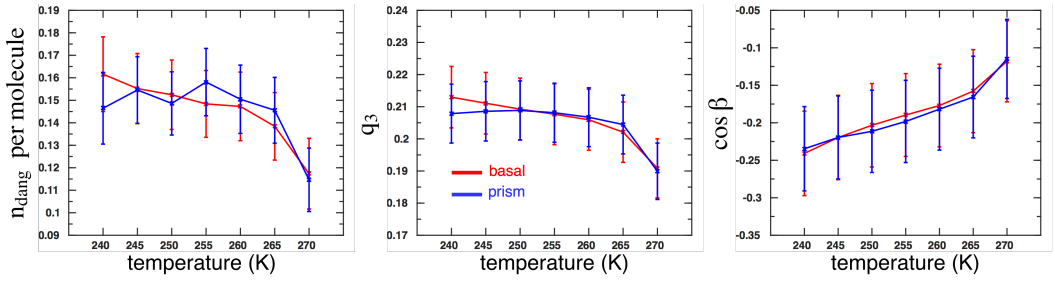


Figure 4.19: Average number of dangling bonds per molecule (left), average q_3 of dangling molecules (centre), and average $\cos\beta$ of dangling molecules (right) in our TIP4P/Ice simulations.

experimental studies of the orientation of dangling bonds on the surface of ice, which measured a weak temperature dependence within the range between 240 K and 265 K [66]. Indeed, our results show a very weak trend towards increased disorder as temperature increases which is more strongly pronounced on the basal surface than the prism surface. At 270 K, the q_3 parameter of the dangling molecules is smaller suggesting the dangling molecules are more liquid-like than at lower temperatures. Given the much larger QLL size at 270 K compared to other temperatures, this result is reasonable. The average $\cos\beta$ indicates that dangling bonds tend to all point towards the surface which is the expected result. As temperature increases, the angle between the dangling bonds and the z-axis increases as the dangling bonds point less directly to the surface. However, between 240 K and 265 K, the average θ angle changes by less than 3° which further endorses the observation that the basal and prism dangling bond structure is nearly constant between 240 K and 265 K.

Dangling bonds occur at the surface. Study of the orientation of dangling bonds therefore provides a commentary on the structure at the surface, however, further study is required to characterise the structure of the entire QLL.

4.4 Ice and QLL Bilayer Structure

Heat maps of the molecular orientation provide a useful measure to help us qualitatively explain the differences in structure of the basal and prism QLL and ice sublayers. Figure 4.20 shows heat maps of the probability distribution of $\cos\beta$ and γ for slices corresponding to peaks in the density profiles at 240 K. Traversing the figure from left to right corresponds to slices or density peaks deeper and deeper into the bulk of the slab. The heat map on the far left corresponds to the outer quasi-liquid sublayer, whilst the final heat map on the right corresponds to the fourth peak in the density profile which at 240 K corresponds to bulk ice. All subsequent heat maps in this section are presented using the same format as Figure 4.20.

Figure 4.21 shows the orientation heat maps for the basal and prism surfaces for all the temperatures. As with Figure 4.20, the far left plots are the outer QLL sublayers

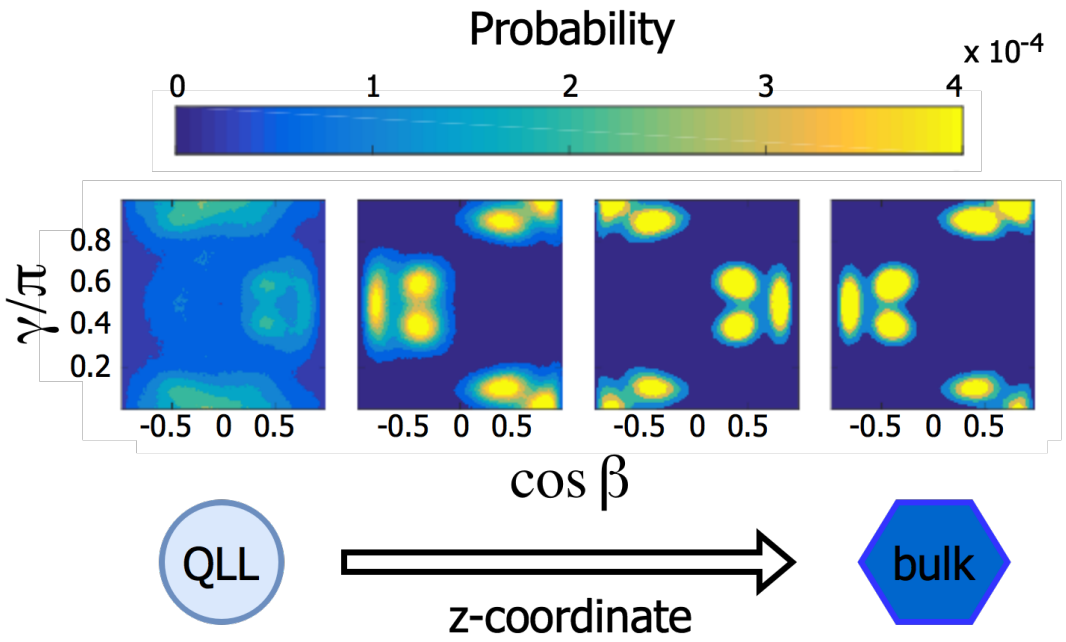


Figure 4.20: Heat maps of water molecule orientations within slices of the basal surface at 240 K using TIP4P/Ice. Each slice corresponds to a peak in the density profile. The first heat map on the left corresponds to the outer sublayer in contact with vacuum and the following three heat maps correspond to slices deeper and deeper into the surface. $\cos\beta$ on the x-axis is the orientation of the molecular dipole, whereas γ on the y-axis is the orientation of the hydrogen-hydrogen vector.

in contact with vacuum. The heat maps clearly illustrate the disorder/order transition which occurs when crossing from the QLL into ice. The basal and prism ice sublayers have distinct orientational structure which is dependent on the crystallographic surface but independent of temperature. The QLL structure is dependent on the crystallographic surface, however, the outer sublayer structure of the surface appears uniform between 240 K and 265 K.

The density heat maps in Figure 4.22 show the 2-D in-plane molecular densities within the basal and prism TIP4P/Ice sublayers. Whilst the basal and prism ice structure is clearly indicated, the disorder/order transition is less well-pronounced compared to the orientation heat maps. Between 240 K and 265 K, the outermost QLL heat maps show an ordered skeletal structure and indicate that the ordered ice structure penetrates into the QLL regions.

The q_3 heat maps in Figure 4.23 show in-plane q_3 distributions within sublayers. Yellow on the scale bar corresponds to ice-like molecules with large q_3 values. Blue corresponds to liquid-like molecules with smaller q_3 values. Traversing from left to

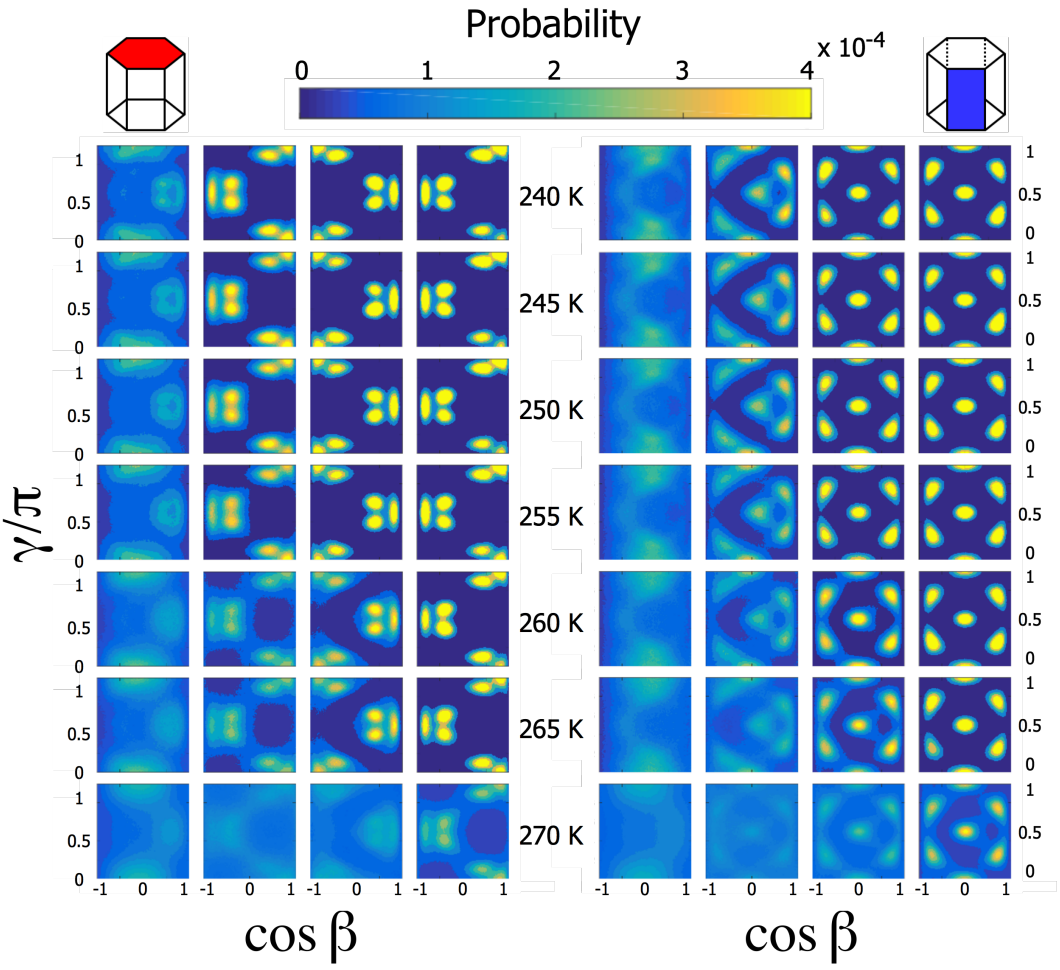


Figure 4.21: Heat maps of water molecule orientations within slices corresponding to density peaks in the basal (left) and prism (right) TIP4P/Ice slabs at 240 K to 270 K. The first heat map on the left corresponds to the outer sublayer in contact with vacuum and the following three heat maps correspond to slices deeper and deeper into the surface. A transition from disorder to order is observed going from the surface into the bulk.

right along the figures, the heat maps become characterised by more and more yellow regions indicating increased crystallinity. The disorder/order transition, however, is less pronounced than in the orientation heat maps.

The dipole heat maps in Figure 4.24 show the in-plane $\cos\beta$ distribution within sublayers. Within the QLL, molecules are not constrained and point in all directions. The average $\cos\beta$ values within the bins, therefore, average out to zero. Within the ice sublayers, molecules are constrained and point in distinct directions. The average $\cos\beta$ values within the bins are, therefore, large positive or negative values as the dipole orientation of neighbouring molecules is correlated. The dipole heat maps

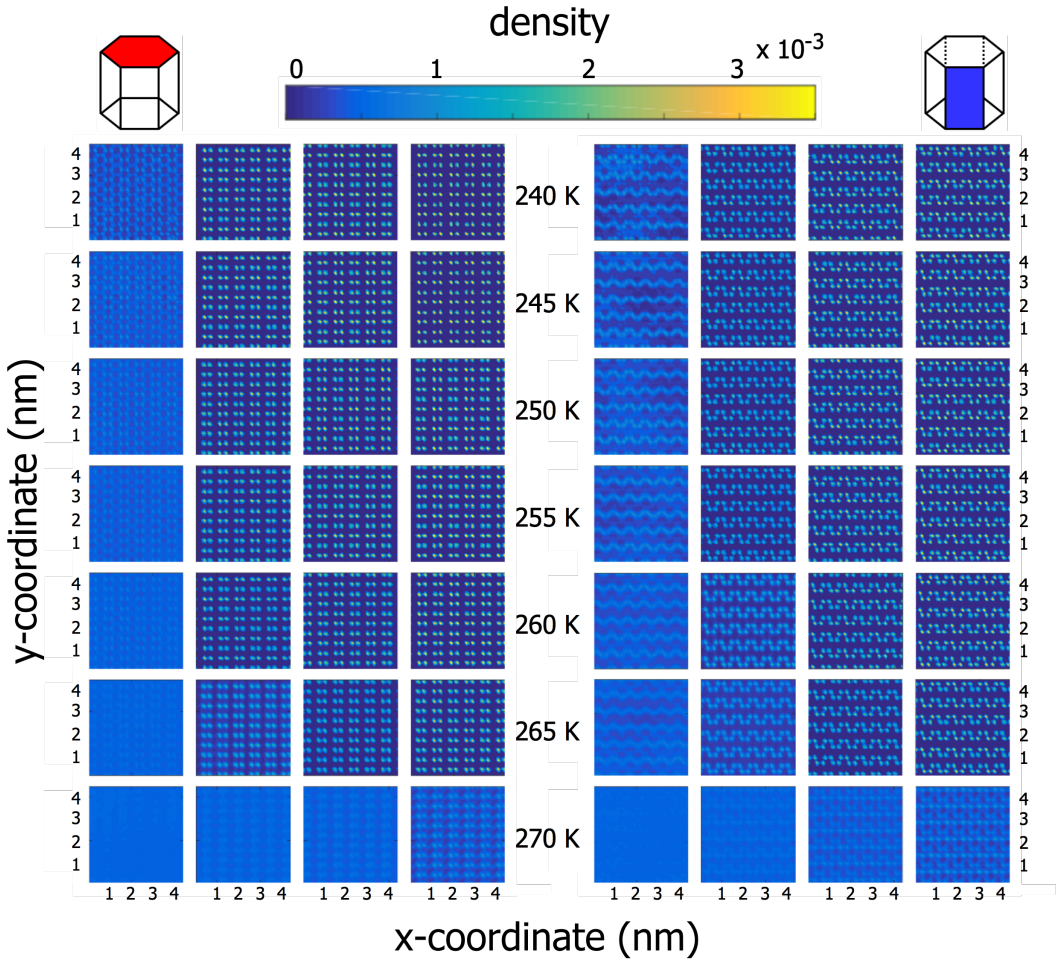


Figure 4.22: Heat maps of the layer-by-layer density density distribution within QLL and ice sublayers calculated from TIP4P/Ice NVT simulations. The x- and y-axes correspond to the two in-plane dimensions of the system. The left column corresponds to the basal slabs and the right column is the prism slabs.

illustrate the increased orientational order of molecules within ice.

The orientation, dipole, q_3 and density heat maps clearly show the presence of a disordered QLL on the surface of ice. The q_3 and density heat maps are less useful for distinguishing clearly the QLL/ice interface and the disorder/ordered regions. The orientation heat maps of $\cos\beta$ and γ and the dipole heat map of $\cos\beta$ show a more pronounced transition from disorder to order. In Chapter 7, the orientation heat maps are used to indicate the position of the QLL/ice interface. The heat maps clearly illustrate the increased disorder of the sublayers with temperature and the expansion of the QLL at 270 K. There are clear structural differences between the QLL and ice and important structural differences between the basal and prism ice

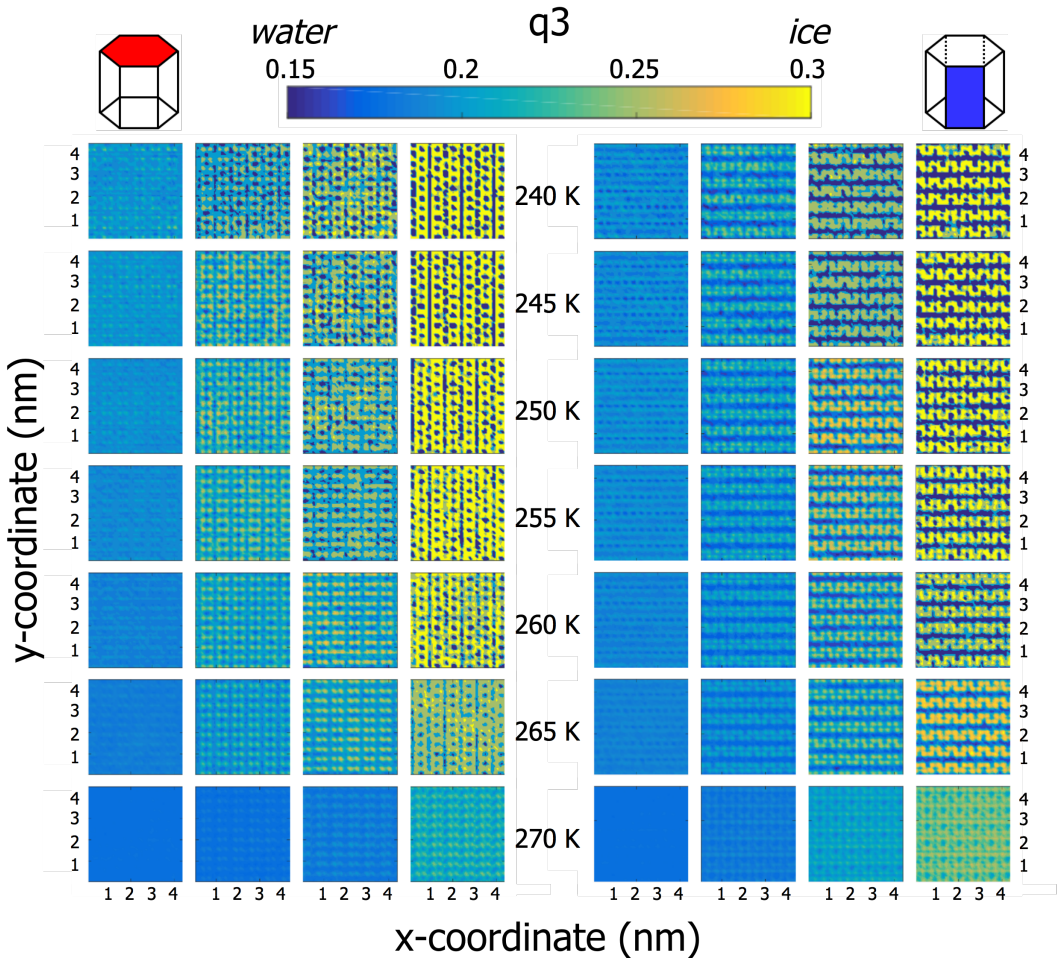


Figure 4.23: Heat maps of the layer-by-layer q_3 distribution within QLL and ice sublayers calculated from TIP4P/Ice NVT simulations. The x- and y-axes correspond to the two in-plane dimensions of the system. The left column is the basal surface and right column is the prism surface.

sublayers. Quantifying how these structural differences influence the rate of growth at the two crystallographic surfaces is key to this project.

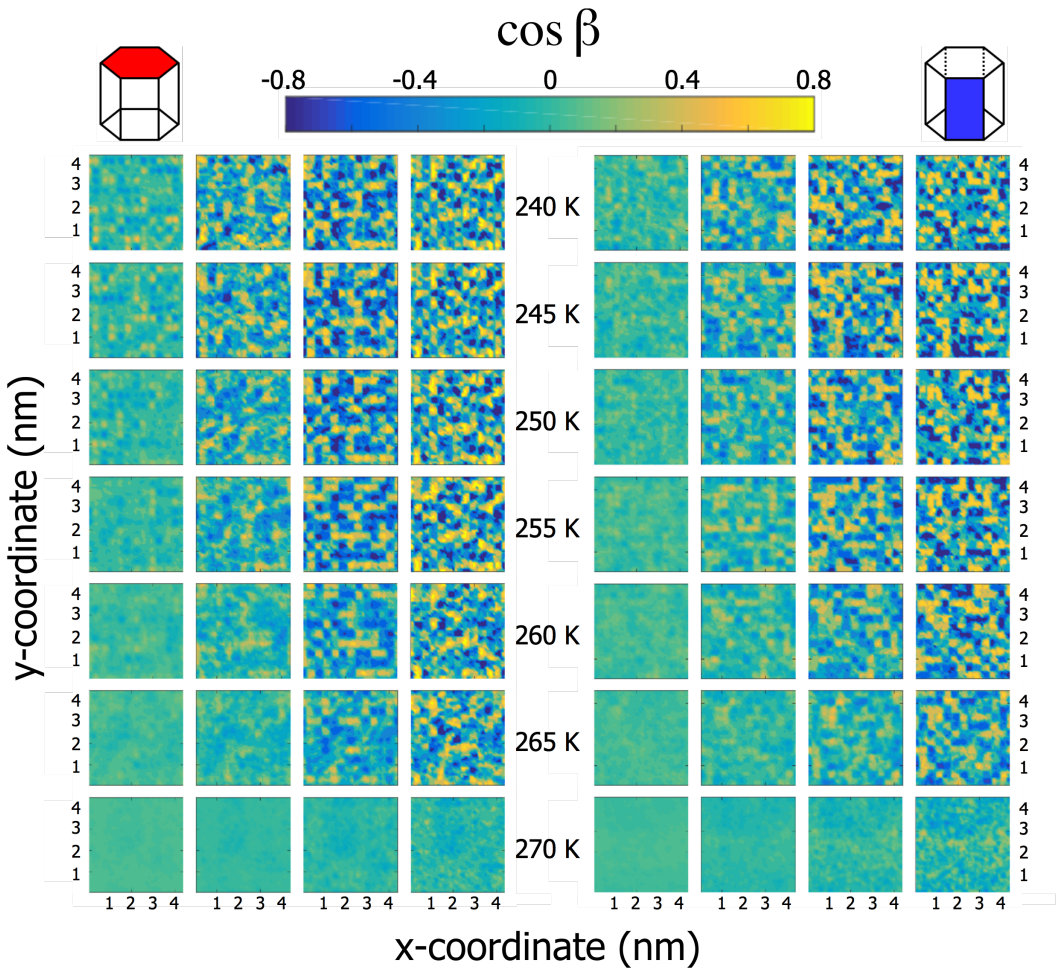


Figure 4.24: Heat maps of the layer-by-layer $\cos\beta$ dipole distribution within QLL and ice sublayers calculated from TIP4P/Ice NVT simulations. The x- and y-axes correspond to the two in-plane dimensions of the system.

4.5 Dynamics of the Quasi-Liquid-Layer

4.5.1 Translational Diffusion

The mobility of QLL molecules can be described through the mean square displacement (MSD). The one-dimensional MSD can be expressed as

$$MSD(t) = \langle(\mathbf{R} - \mathbf{R_0})^2\rangle = \frac{1}{N}\sum_{i=1}^{N}(R_i(t) - R_i(0))^2 \tag{4.5}$$

where N is the number of particles, $R_i(0)$ and $\mathbf{R_0}$ are the reference coordinate of each particle and $R_i(t)$ is the coordinate of particle i at time t. In order to take into account any drift in the centre of mass during molecular dynamics simulations the

following modified form of the MSD is used

$$MSD(t) = \frac{1}{N}\sum_{i=1}^{N}(R_i(t) - R_i(0) - [R_{cm}(t) - R_{cm}(0)])^2 \tag{4.6}$$

where $R_{cm}(t)$ is the centre of mass of the system at time t. Contrasting our results when the centre of mass is and isn't taken into account shows that no drift occurs during our simulations.

The one-dimensional MSD is calculated in the two in-plane directions using Eq. 4.6 for the oxygen atoms in our TIP4P/Ice and mW simulations. To perform the analysis, the NVT simulations are split into two equal halves. The first half is used to generate $\sim$ 300 equally spaced reference frames. For each reference frame, the MSD is calculated for a fixed length of time equal to half the simulation length. For a typical MD simulation in this work, the first 10 ns are discarded for annealing and the remaining 90 ns are split into two and the MSD is calculated for 45 ns for each reference frame. The coordinates are unwrapped to take into account the increasing

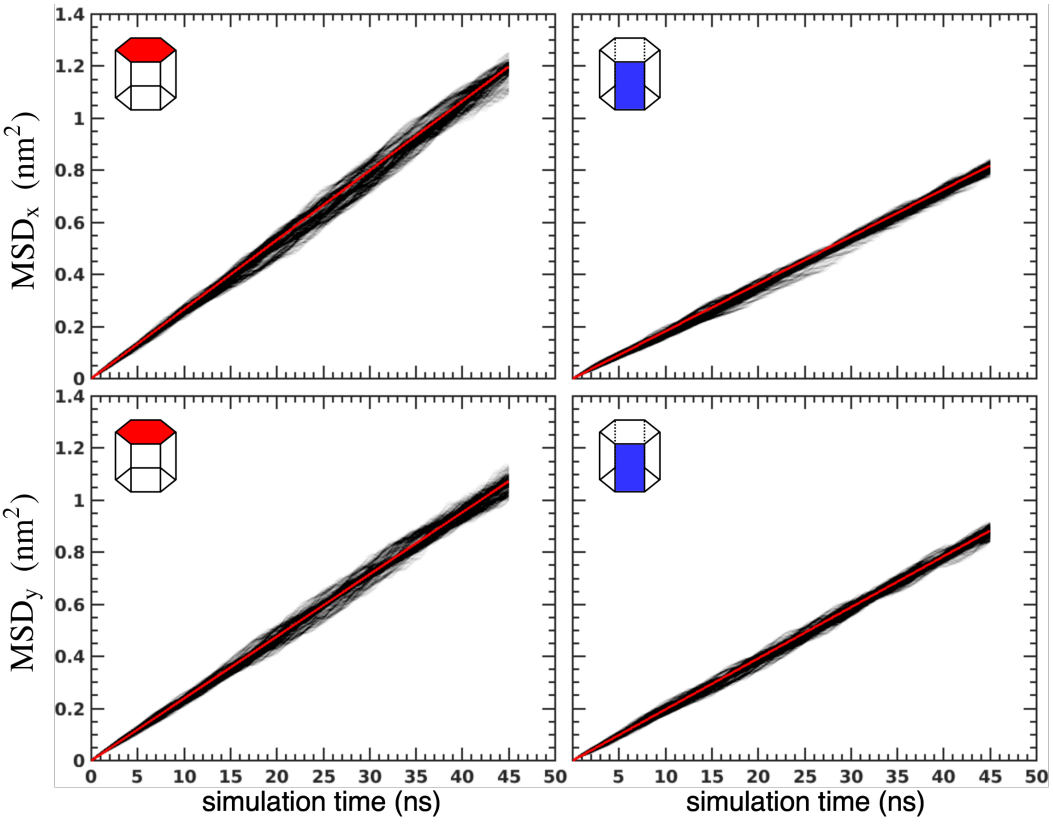


Figure 4.25: 1-D MSD of the basal (left) and prism (right) TIP4P/Ice molecules at 255 K.

separation of molecules as they cross the box boundary. Figures 4.25 and 4.26 show the $\sim$ 300 plots of MSD(t) for the TIP4P/Ice and mW surfaces, respectively, at 255 K. Similar results are obtained for the two models at the other simulated temperatures. The trend is linear in all cases, a characteristic result for the MSD of liquids. Red lines in the plots correspond to linear fits to the data.

The one-dimensional diffusion coefficient, D can be obtained using Einstein's relation from the slope of the one-dimensional MSD [116]

$$D = \frac{1}{2}\frac{\mathrm{d}MSD}{\mathrm{d}t} \tag{4.7}$$

Bulk molecules are immobile on the timescale of our simulations and do not contribute to the MSD. The surface diffusion coefficient, D^* is defined as

$$D^* = D/Q \tag{4.8}$$

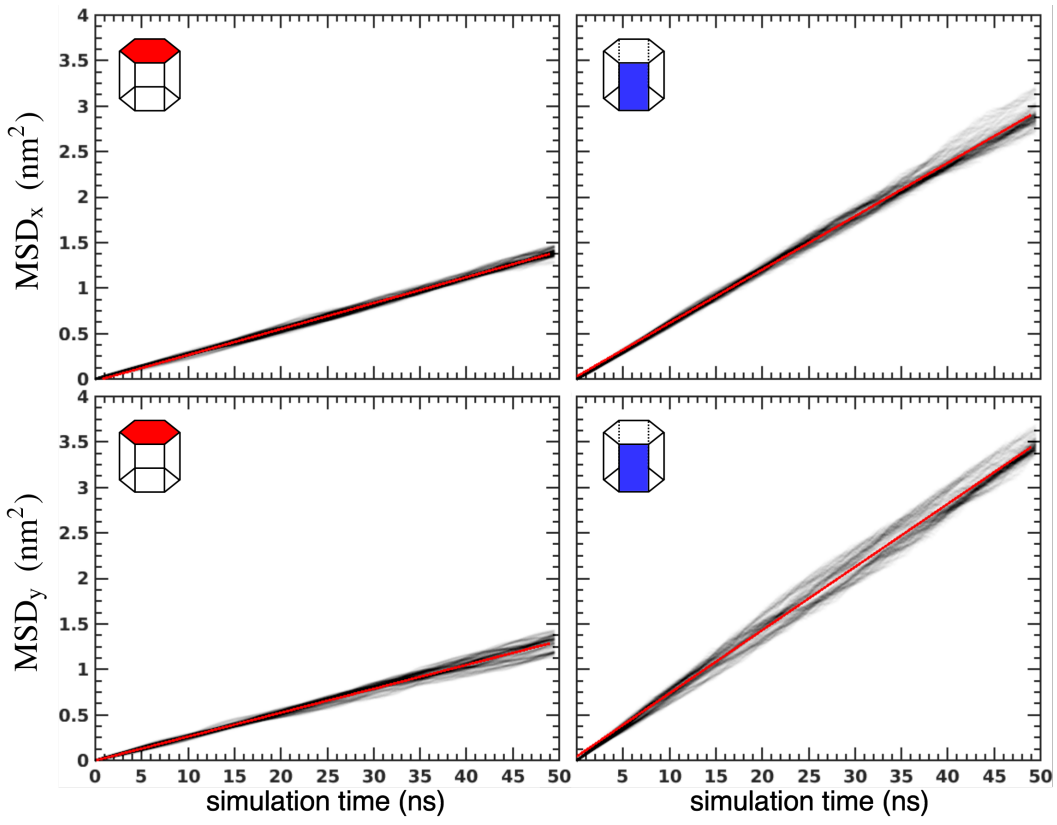


Figure 4.26: 1-D MSD of the basal (left) and prism (right) mW molecules at 255 K.

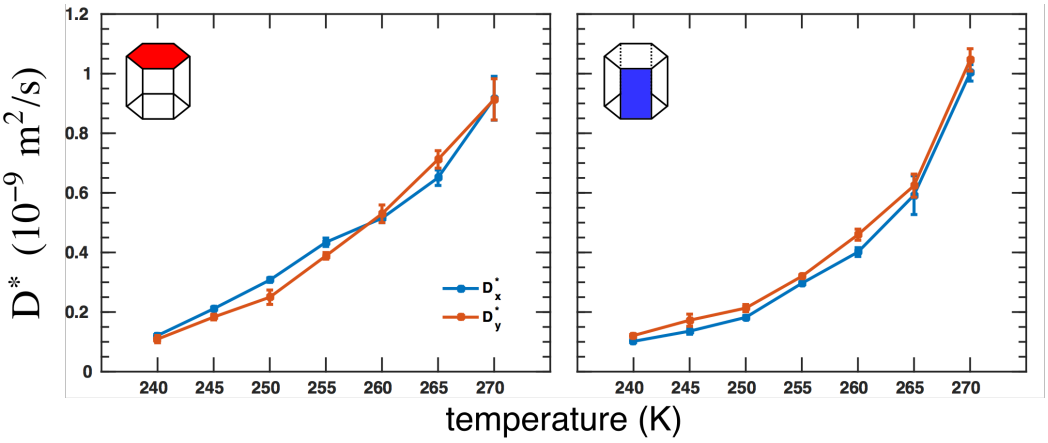


Figure 4.27: Translational one-dimensional in-plane diffusion coefficient of basal (left) and prism (right) TIP4P/Ice molecules.

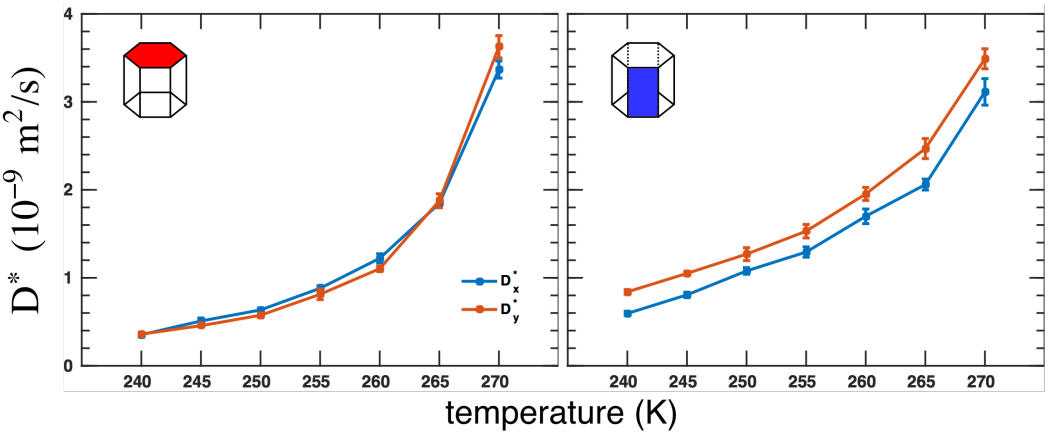


Figure 4.28: Translational one-dimensional in-plane diffusion coefficient of basal (left) and prism (right) mW molecules.

where $Q = N_{liquid}/N_{total}$ is the average proportion of liquid-like molecules in the system. Figures 4.27 and 4.28 show D^* calculated for diffusion in the two in-plane directions based on linear fits to MSD(t) for the TIP4P/Ice and mW surface, respectively. The error bars indicate the standard deviation of the fits obtained from the many reference frames. Arrhenius plots of the logarithm of D^* against the inverse of temperature are shown in Figures 4.29 and 4.30 for TIP4P/Ice and mW, respectively. Within the error range, the trends are near linear suggesting diffusion can be considered an activated process which obeys Boltzmann statistics. The functional form of D^* can broadly be considered to be of the form $D^* \propto \exp\left(\frac{-E_a}{k_B T}\right)$, where E_a is the activation energy of diffusion. An Arrhenius fit to our results yields activation energies of 36.09 ± 4.16 kJ/mol and 39.16 ± 5.24 kJ/mol for the basal

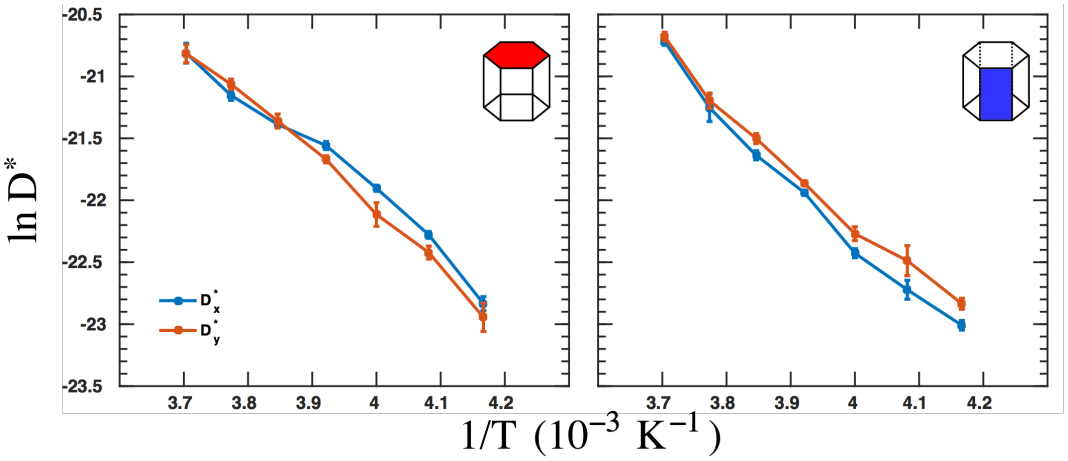


Figure 4.29: Logarithm of the one-dimensional diffusion coefficients of the basal (left) and prism (right) TIP4P/Ice surfaces plotted against the inverse of temperature.

and prism TIP4P/Ice surfaces, respectively. Arrhenius fits to our mW results yields activation energies of 39.06 ± 9.19 kJ/mol and 26.32 ± 3.32 kJ/mol. The uncertainty correspond to the 95% confidence interval. The TIP4P/Ice barrier heights are similar to the energy required to break two H-bonds, measured experimentally at 46 kJ/mol [117], which supports the notion that in-plane diffusion involves breaking two hydrogen bonds. Previous MD simulation works using a six-site water model yield an activation energy for translational diffusion increasing from 29.1 kJ/mol at around $-60°C$ to 53.8 kJ/mol close to the melting point [47, 48]. A previous study using TIP4P/2005 reports a diffusion activation energy of around 10 kJ/mol at around $-60°C$, increasing to around 45 kJ/mol close to the melting point [47]. These results are in agreement with our MD simulation estimates. Simulations using TIP4P estimate a low temperature activation energy of 23.2 kJ/mol between 180 K and 210 K [45].

A two-dimensional surface diffusion coefficient can be calculated from the one-dimensional surface diffusivities as [47, 48]

$$D^*_{i,j} = (D^*_i + D^*_j)/2 \tag{4.9}$$

Figure 4.31 shows the two-dimensional surface diffusion coefficient, defined in Eq. 4.9, for the TIP4P/Ice and mW surfaces. The error bars indicate the standard deviation

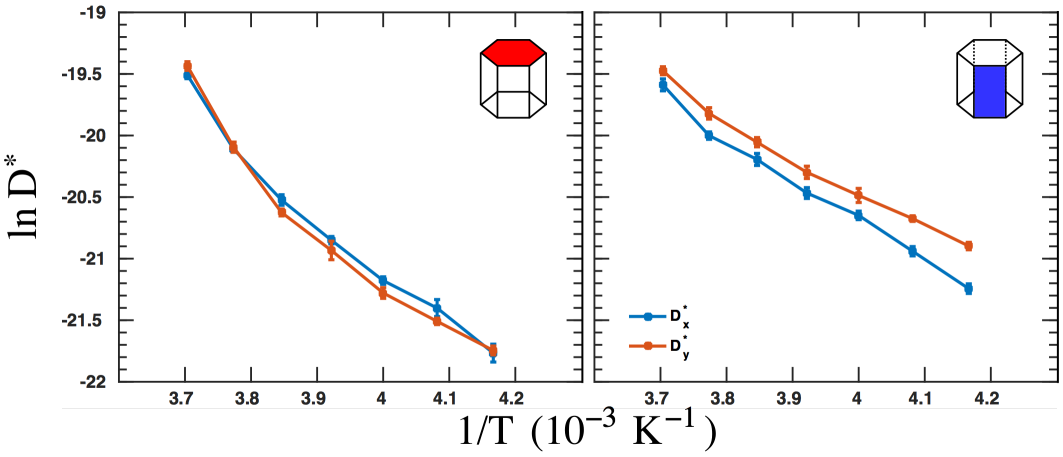


Figure 4.30: Logarithm of the one-dimensional diffusion coefficients of the basal (left) and prism (right) mW surfaces plotted against the inverse of temperature.

across the many reference frames. The results show that in-plane translational diffusion increases with temperature. The range of our simulation estimates of the translational diffusion coefficient using TIP4P/Ice quantitatively agree with results using the 6-site NE6 water model [48, 118]. NMR experimental measurements of the translational diffusion coefficient of supercooled water yield a value of 0.650×10^{-9} m^2/s at 260.1 K, 0.531×10^{-9} m^2/s at 255.3 K, 0.414×10^{-9} m^2/s at 250.8 K and 0.204×10^{-9} m^2/s at 239.8 K [43]. These measurements are slightly larger than our estimates, however, they are nonetheless in agreement with our results because they are measured using supercooled water which has a greater mobility than the QLL. Our findings can be further validated by comparing our results with experimental

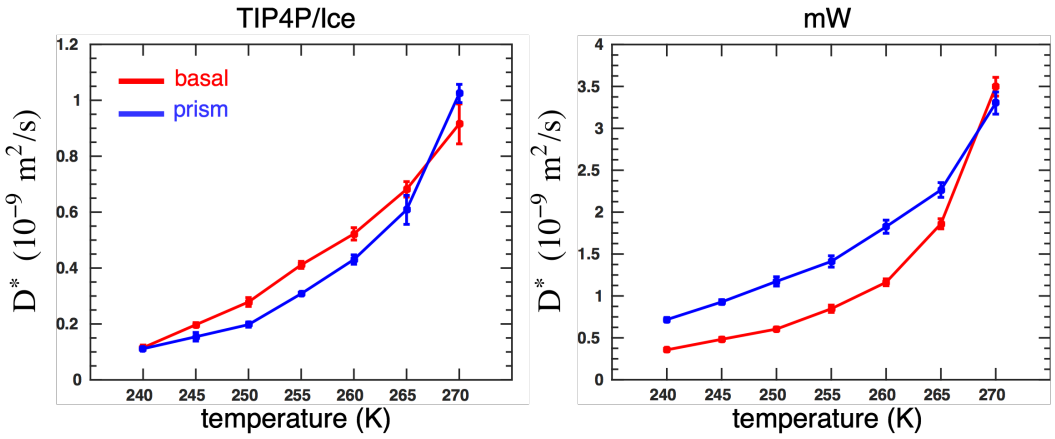


Figure 4.31: Two-dimensional translation diffusion coefficient of for TIP4P/Ice (left) and mW (right)

measurements of the self-diffusion coefficient on ice surfaces obtained by studying the formation of grain boundaries [119]. At 263.15 K a value of 0.3×10^{-9} m^2/s is measured and at 258.15 K a value of 0.29×10^{-9} m^2/s, in close agreement with our TIP4P/Ice results.

Whilst D_x^* depends on Q, the ratio of the one-dimensional diffusivities does not, and $\frac{D_x^*}{D_y^*} = \frac{D_x}{D_y}$. Figure 4.32 shows the anisotropy of the in-plane translational diffusion of our TIP4P/Ice and mW simulations. A value of one is perfect isotropy. Our result shows that the diffusion is isotropic at all the simulated temperatures within reasonable margin. However, as temperature decreases diffusion becomes less isotropic and tends towards anisotropy at low temperatures. This is consistent with previous simulation results [47, 48] and provides further validation of our results.

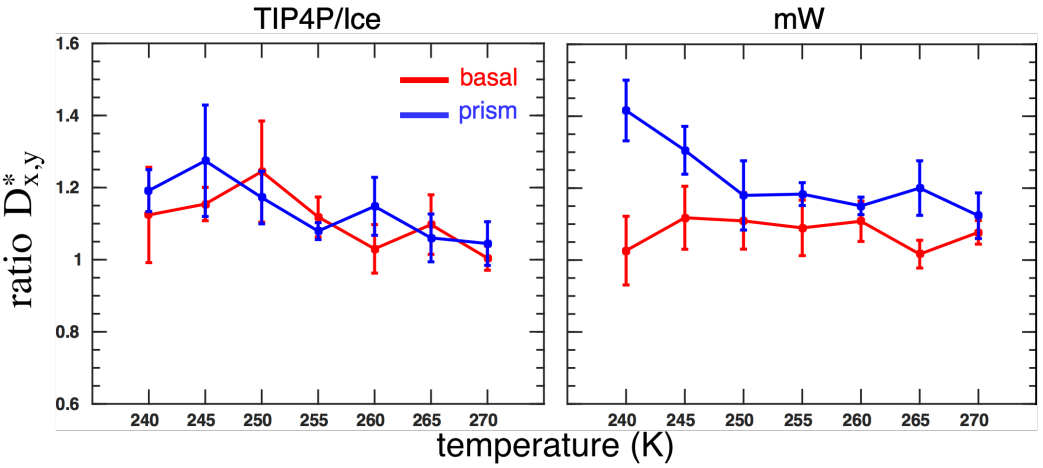


Figure 4.32: Ratio of the one-dimensional in-plane translation diffusion coefficients of the TIP4P/Ice (left) and mW (right) surfaces. Large values indicates anisotropic in-plane diffusion whilst values close to one indicate isotropic diffusion.

4.5.2 Rotational Diffusion

The rotational diffusion coefficient, D_{rot} can be calculated using the same approach as the translational diffusion coefficient. The MSD of the dipole orientation angle, β is calculated and plotted in Figure 4.33 at 245 K and 265 K using TIP4P/Ice. The result is equivalent to the translational MSD result and shows a linear increase with simulation time. The red line indicates the best fit to the plots. The orientation of water molecules is not captured by the coarse-grained mW model which does not

include hydrogen atoms. As a result, it is not possible to determine the rotational diffusion from the mW model using this approach.

The rotational diffusion coefficient can be calculated from the gradient of MSD(t) using the same approach as the translational diffusion coefficient. Figure 4.34 is the TIP4P/Ice rotational diffusion coefficient plotted against temperature. The Figure also shows the plot of $\ln(D^*)$ plotted against $1/T$. The trend is similar to the translational diffusion trend in Figure 4.31. The rotational diffusion coefficient increases with temperature and displays an exponential-like functional form, characteristic of a rotational energy barrier which can be overcome with a Boltzmann probability. The numerical estimates of the rotational diffusion coefficients are of the same order as the translational diffusion coefficients. An Arrhenius fit to the TIP4P/Ice rotational diffusion coefficients yields an activation energy of 14.12 ± 1.60 kJ/mol and 16.84 ± 5.14 kJ/mol for the basal and prism surfaces, respectively, with 95% confidence. The rotation energy barriers are lower than the translation barriers.

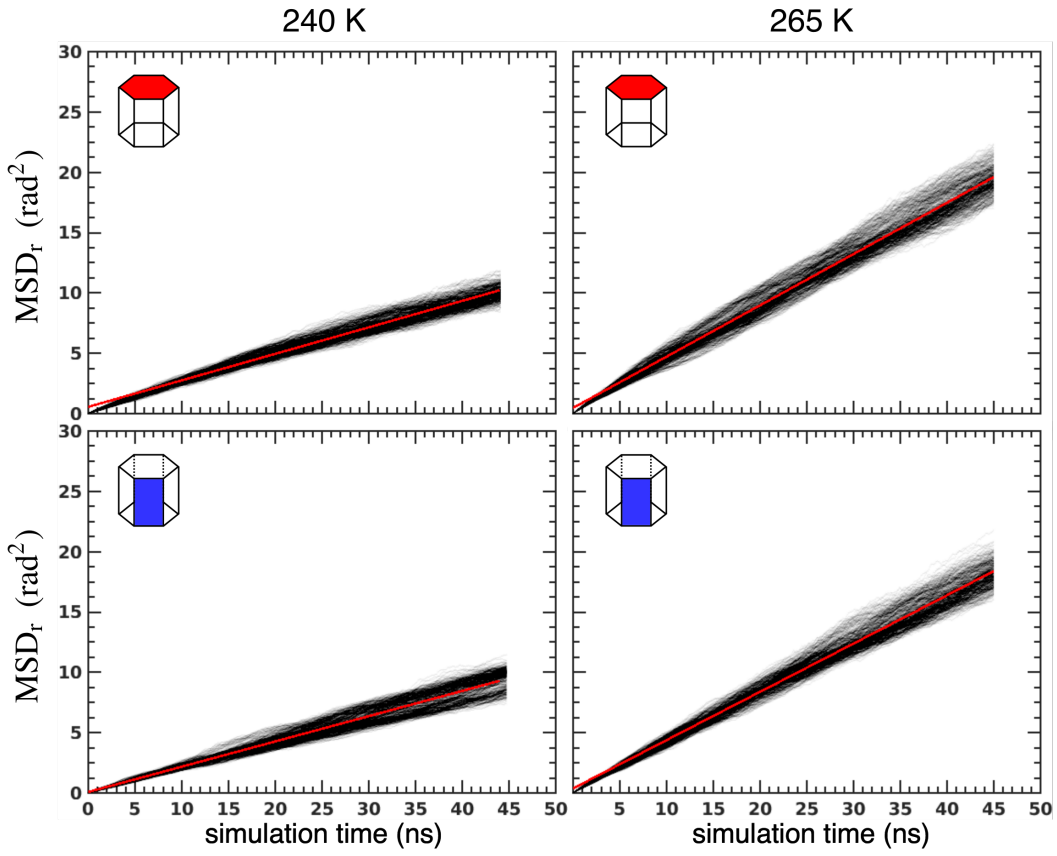


Figure 4.33: Rotational $\cos\beta$ MSD of TIP4P/Ice simulations at 240 K (left) and 265 K (right).

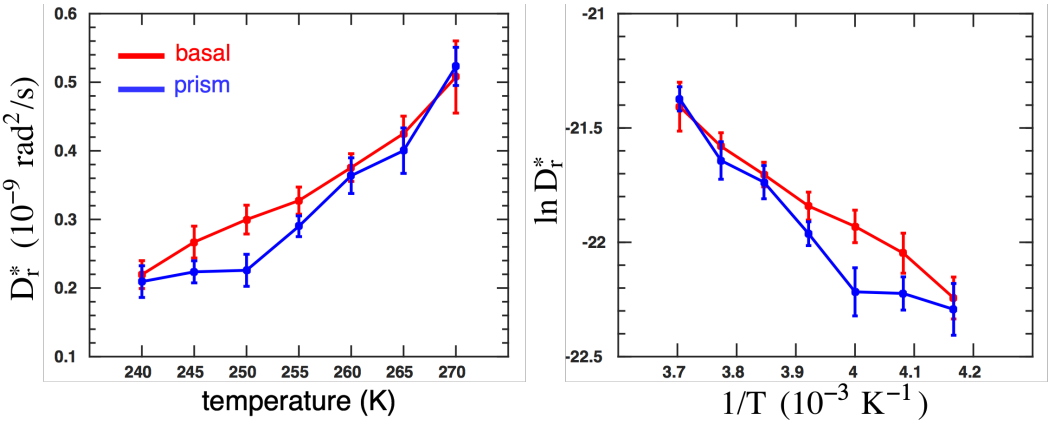


Figure 4.34: Rotational diffusion coefficient against temperature (left) and its Arrhenius plot (right) for TIP4P/Ice

This makes intuitive sense because translation entails a greater degree of reorientation of molecules and involves breaking more hydrogen bonds than rotating. Simulation studies of supercooled water and normal water using TIP4P/2005 between −40°C and +60°C yields a rotation activation energy of 13.8 kJ/mol [120], in good agreement with our simulation estimates.

4.6 Conclusions

Our simulation results show that the QLL size increases with temperature. With TIP4P/Ice the QLL size obeys a logarithmic relationship, in agreement with previous experimental and simulation results [1, 7, 29, 33–38]. Close to the melting point at 270 K, analysis of the TIP4P/Ice liquid-like clusters suggest that the QLL is melting and crystallising at least two bilayers. mW, however, is over-structured and describes the QLL thickness and dynamics less accurately than TIP4P/Ice.

Analysis of the average dipole moments of TIP4P/Ice basal and prism QLL molecules suggests that between 240 K and 265 K, the local environment changes very little and is fairly uniform across the two crystallographic surfaces. The orientation of dangling bonds does not display any strong dependence on the crystallographic surface and remains uniform between 240 K and 265 K. These observations are consistent with previous vibrational response experiments [66, 69]. The implication

of our studies is that the outer QLL structure at the interface with vacuum or vapour depends only weakly on the underlying crystalline lattice.

From a theoretical perspective, the equilibrium QLL size can be modelled by describing the atoms as vibrating particles in a harmonic potential. Modelling the intermolecular forces using an exponentially decaying function yields the logarithmic relationship of the QLL with temperature, as observed in simulation and experiments. An Arrhenius fit to the our translation diffusion coefficients yields activation energies which indicate that diffusion involves breaking two H-bonds. Previous simulation works report that molecules within the outer QLL sublayer at the interface form on average two hydrogen bonds [29]. Our activation energies are consistent with a mechanism of diffusion involving the breaking of hydrogen bonds and agree with previous simulation estimates using a range of water models [45, 47, 48].

Contrasting our TIP4P/Ice and mW translational self-diffusion coefficient estimates with experimental measurements reveals that TIP4P/Ice successfully describes the dynamics of the QLL. mW overestimates the coefficients and has much faster intrinsic dynamics. We can infer from these observations that TIP4P/Ice simulations will yield insightful predictions of the QLL kinetics whilst mW may not. As temperature increases, the QLL grows in thickness and becomes more mobile. Our diffusion results show that the QLL has distinct liquid-like kinetics properties. Metadynamics simulations can be used to understand the thermodynamics of the QLL.

Chapter 5

Energetics of the Quasi-Liquid-Layer from Metadynamics Simulations

The free energy associated with ice growth from the QLL can be estimated using well-tempered metadynamics. Metadynamics is an enhanced sampling technique designed to reconstruct coarse-grained free energy surfaces as a function of one or more slowly varying degrees of freedom known as collective variables (CVs). Our CV in the metadynamics simulations described in this Section is the number QLL molecules, n_l, calculated using q_3. Details of the simulation protocol are described in Section 3.5.1. Our metadynamics simulations bias only one selected surface within each slab. Biasing the number of QLL molecules enables us to the recover the free energy of the system as a function of the QLL size and uncover the energy cost associated with melting and crystallising the QLL.

Computing n_l involves calculating q_3 which is computationally expensive. In order to offset the heavy cost involved in calculating n_l, we perform metadynamics simulations using mW which is 2 orders of magnitude faster than atomistic models such as TIP4P/Ice [94]. In Chapter 4 we noted that mW is poor at capturing the kinetics of the QLL, however, previous studies have shown that mW is good at describing the energetics and thermodynamic anomalies, and therefore is a suitable

choice for computing free energies using metadynamics [94]. In this work we perform metadynamics simulations of the mW basal and prism surfaces at 240 K, 260 K, 265 K and 270 K. Simulations are also performed using TIP4P/Ice at 260 K to show the effect of the water model on the results.

5.1 mW Metadynamics

Figure 5.1 shows n_l during the mW metadynamics simulations at 240 K, 260 K, 265 K and 270 K. Whilst the QLL thickness fluctuates around an equilibrium value during MD simulations, with metadynamics n_l is biased resulting in large fluctuations and frequent melting and crystallisation events. The sampling of QLL melting and crystallisation is enhanced and the underlying free energies can be reconstructed. Simulations are performed until the free energies are reasonably converged. The convergence of the free energies is described in detail in Section 5.3. In practice the simulation time required to reach convergence depends on both the temperature

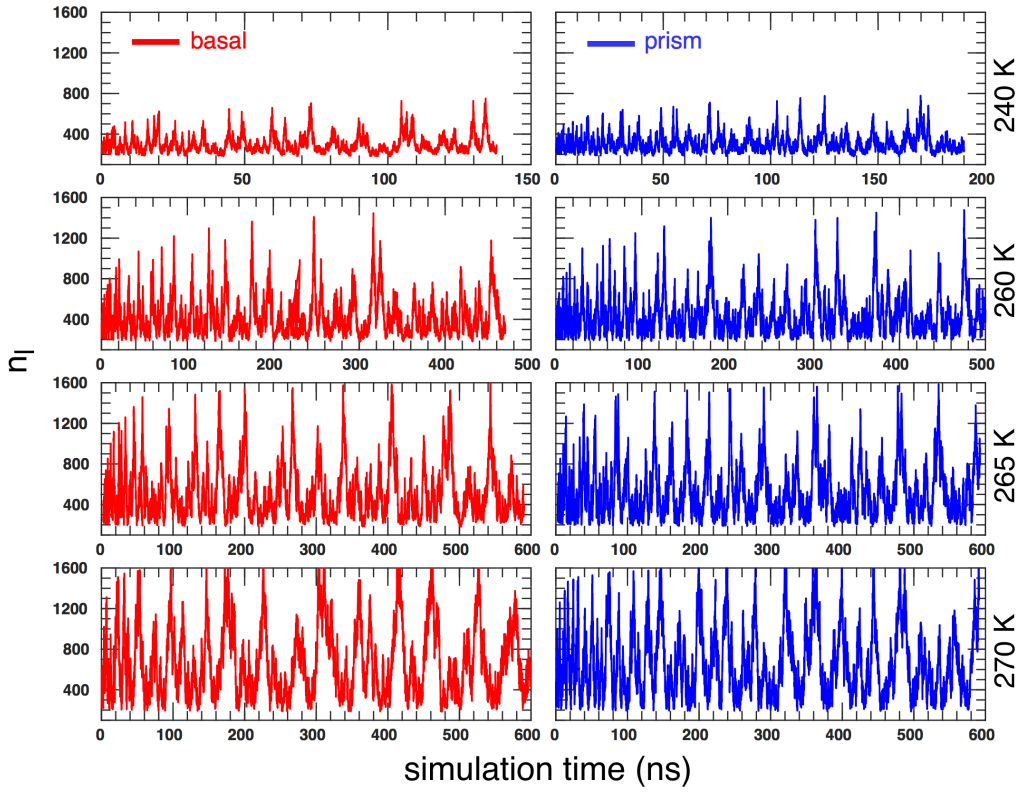


Figure 5.1: Number of liquid molecules, n_l during the mW metadynamics simulations at 240 K, 260 K, 265 K and 270 K.

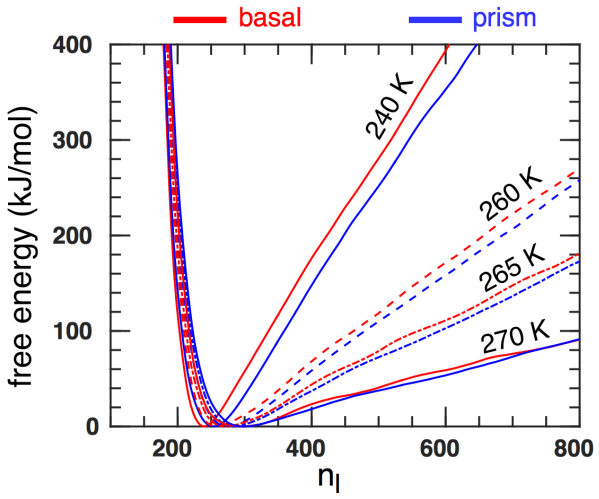


Figure 5.2: Free energy associated with QLL melting and crystallisation calculated from metadynamics simulations at 240 K, 260 K, 265 K and 270 K using mW.

and the water model used. At 240 K the CV explores a narrow range of values, and shorter simulations of up to around 200 ns are performed using mW. At higher temperatures, the system has more thermal energy and explores a broader range of n_l values, spending less time within each state. Simulations of up 600 ns are therefore performed at 260 K, 265 K and 270 K using mW to ensure that the CV space is well sampled.

The free energy profiles associated with crystallising and melting the QLL are reconstructed from our mW simulations 100 times at equally spaced intervals during the last 10 ns of the simulations at 240 K and during the last 20 ns of the simulations at the 260 K to 270 K. A mediated profile is constructed from the 100 reconstructions by averaging the profiles to take into account fluctuations in the free energy at convergence. The mediated free energy profiles are plotted in Figure 5.2 for the mW simulations at 240 K, 260 K, 265 K and 270 K. The results show that there is an energy minimum corresponding to the equilibrium QLL thickness. Melting the QLL further is an unfavourable process which costs energy, whereas crystallising the QLL when it is larger than equilibrium, is a favourable process which does not require overcoming any energy barriers. The free energy profiles include only one minimum corresponding to the equilibrium QLL size. There are no additional local minima corresponding to the melting of additional crystal layers and we can infer that QLL

crystallisation occurs spontaneously when the QLL size is larger than its equilibrium size. The gradient of the free energy profile decreases with temperature indicating that the driving force for crystallising the QLL decreases with temperature. This result is expected given that below the melting point the thermodynamic stability of the liquid phase increases with temperature. The free energy sharply increases for CV values below the free energy minimum. This result is an artifact of the simulation setup as n_l computed using PLUMED has a minimum residual value due to the incorrect assignment of ice molecules at the vacuum interface as liquid. The outcome is that the CV is inherently prevented from exploring values smaller than the number of surface molecules resulting a sharp increase in the free energy for low CV values.

5.2 TIP4P/Ice Metadynamics

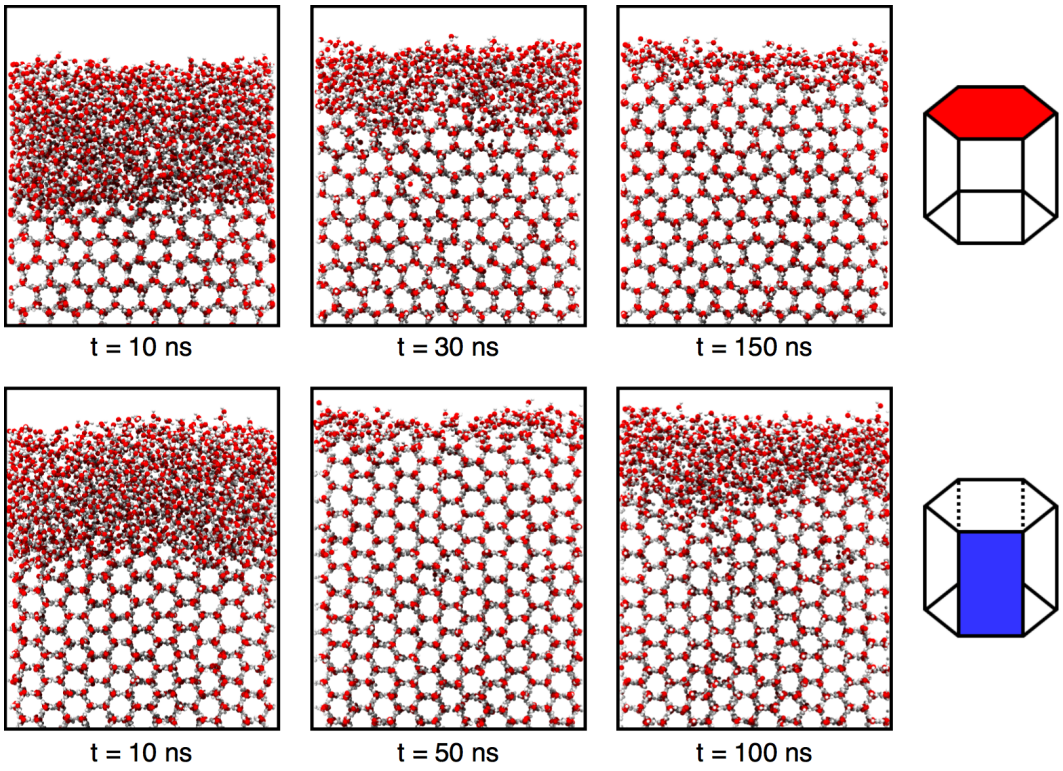


Figure 5.3: Snapshots of the basal and prism TIP4P/Ice metadynamics simulations at 260 K. n_l is biased on the top surface enabling the free energy as a function of n_l to be recovered.

Metadynamics simulations are performed using TIP4P/Ice at 260 K. The simulations are significantly more expensive than mW due to the reduced time step and

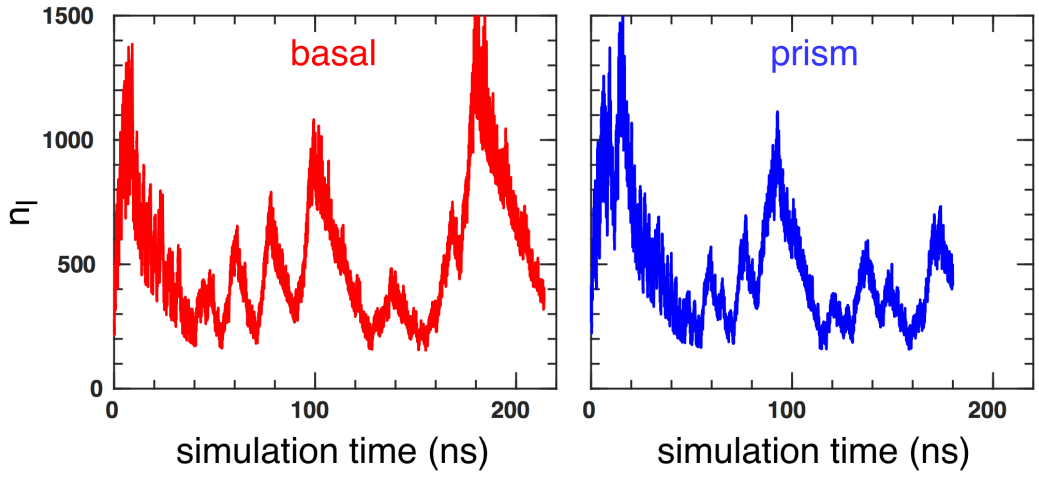


Figure 5.4: Number of liquid molecules, n_l during TIP4P/Ice metadynamics simulations at 260 K

slower dynamics of the QLL. As a result, more computer time is required to simulate less simulation time. Snapshots of the TIP4P/Ice metadynamics simulations are shown in Figure 5.3 and clearly illustrate that the melting and crystallisation of the QLL is sampled during the simulations. Figure 5.4 shows the CV, n_l. As with mW, the plots show many melting and crystallising events of the QLL which do not occur within the time scales of MD simulations.

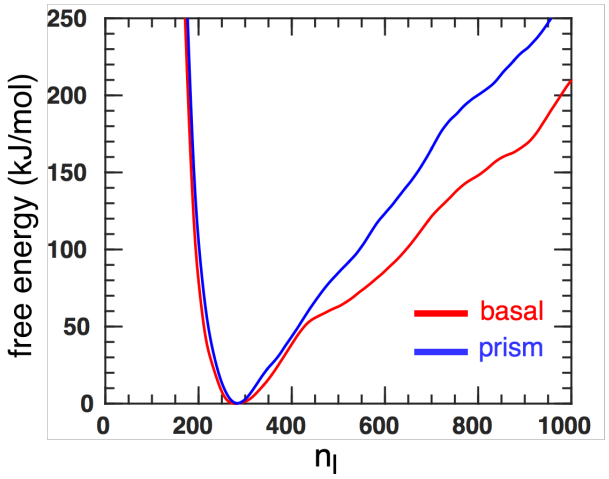


Figure 5.5: Free energy profile associated with QLL melting and crystallisation calculated from our metadynamics simulations using TIP4P/Ice at 260 K.

The free energies associated with crystallising and melting the QLL are reconstructed from our TIP4P/Ice simulations and plotted in Figure 5.5. The profiles are mediated over the last 20 ns of the simulations using 500 equally spaced free energy

profiles. The result is consistent with the mW simulations: There is a free energy minimum corresponding to the equilibrium QLL, and melting the QLL further is an activated process which encounters no other local minima.

5.3 Convergence

In the well-tempered metadynamics framework, the height of the deposited Gaussians decrease with time. Figure 5.6 shows the height of the deposited Gaussians plotted against simulation time using mW. It is clear that the height decreases with time, however, large Gaussians may still be deposited as the height depends on the number of times the current CV value has been visited. As the simulation time and amount of sampling increase, Gaussians with a large height are deposited less frequently. The deposited Gaussian heights during the TIP4P/Ice simulations at 260 K are plotted in Figure 5.7.

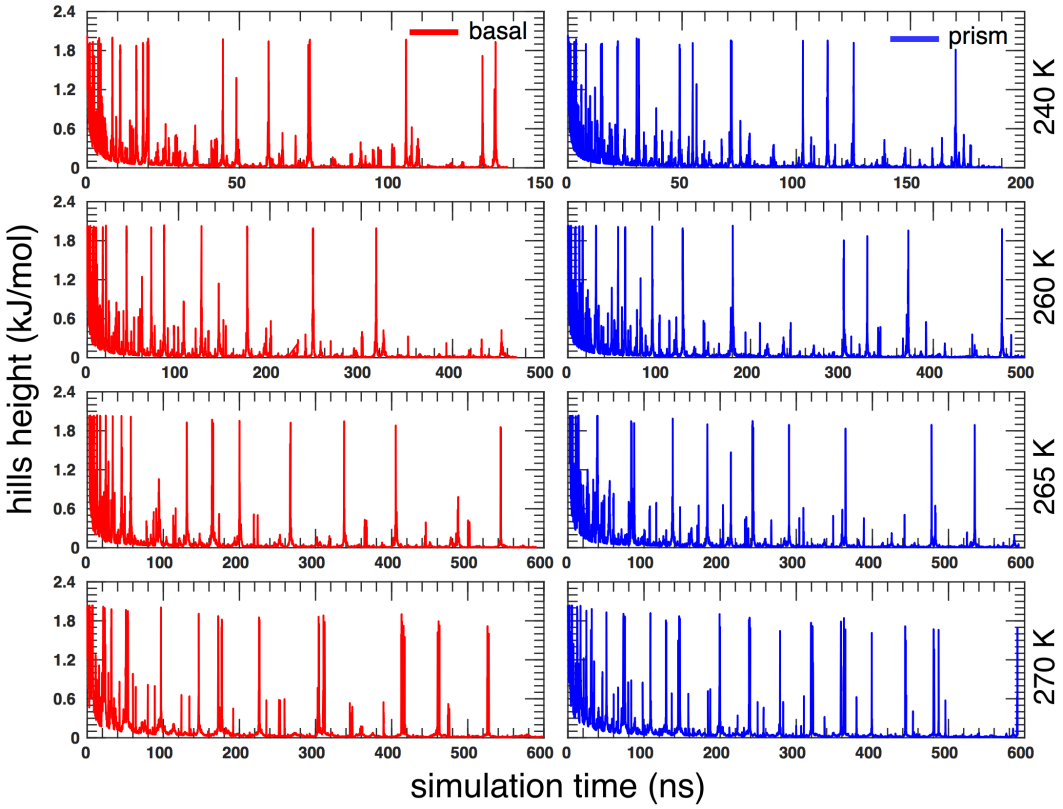


Figure 5.6: Height of deposited Gaussian hills during mW metadynamics simulations at 240 K, 260 K, 265 K and 270 K.

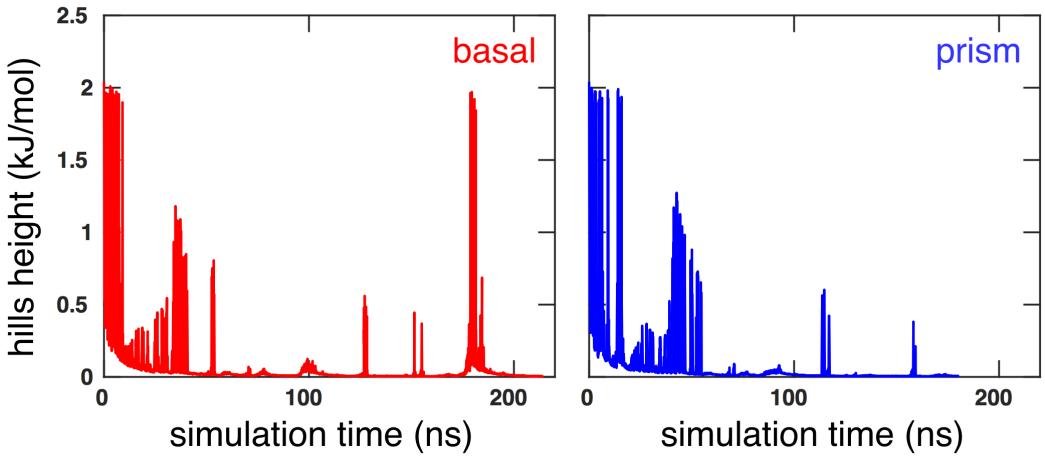


Figure 5.7: Height of deposited Gaussian hills during TIP4P/Ice metadynamics simulations at 260 K

The accuracy of our metadynamics free energy profiles can be quantified by measuring the convergence of the results. The convergence is determined by monitoring $\Delta F(\mathrm{t})$, the difference between the free energy at two set points in the CV space with time. $\Delta F(\mathrm{t})$ is defined as

$$\Delta F(t) = F(n_{l_i})(t) - F(n_{l_0})(t) \tag{5.1}$$

where $F(n_{l_0})$ corresponds to the free energy close to the free energy minimum, and $F(n_{l_i})$ is the free energy at a point n_{l_i} along the CV space. $\Delta F(\mathrm{t})$ is calculated by repeatedly reconstructing the free energy profile during the simulations every time 100 Gaussian hills have been deposited.

The convergence of the mW metadynamics simulations are plotted in Figure 5.8 for the basal and prism simulations at 240 K, and Figures 5.9 and 5.10 for the basal and prism simulations, respectively, at 260 K, 265 K and 270 K. The convergence is calculated using five different values of n_{l_i} and one fixed value of n_{l_0}. The figures also include the final free energy profiles constructed at the end of the simulations and the positions of $F(n_{l_0})$ and the five positions of $F(n_{l_i})$ used to calculate $\Delta F(\mathrm{t})$ are indicated. Large fluctuations in $\Delta F(\mathrm{t})$ indicate that the free energies are not well converged whereas small fluctuations indicate that they are well converged. Our mW convergence results show small fluctuations. As $n_{l_{n_i}}$ increases, the sampling is

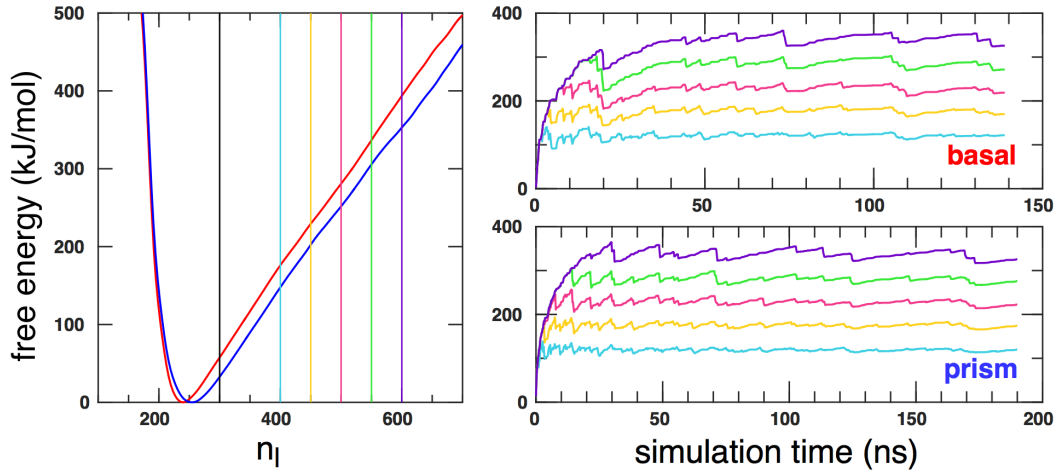


Figure 5.8: Convergence of the basal and prism mW metadynamics simulations at 240 K.

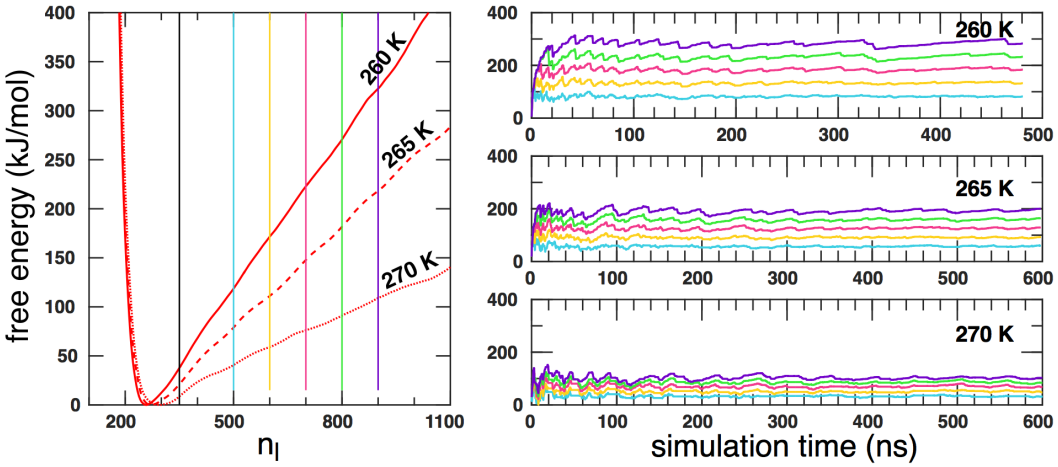


Figure 5.9: Convergence of the basal mW metadynamics simulations at 260 K, 265 K and 270 K.

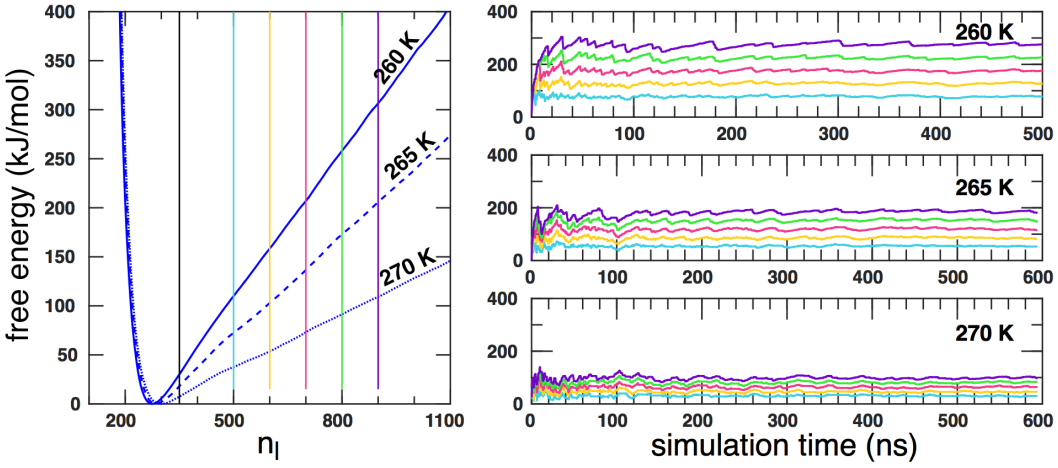


Figure 5.10: Convergence of the prism mW metadynamics simulations at 260 K, 265 K and 270 K.

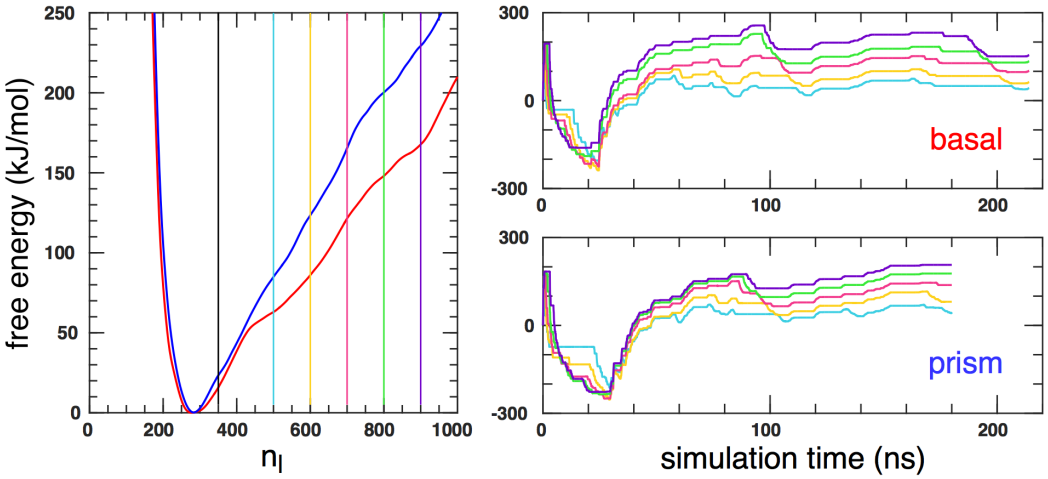


Figure 5.11: Convergence of the basal and prism TIP4P/Ice metadynamics simulations at 260 K.

less well converged because the system tends to sample close to the equilibrium QLL thickness and explores more extreme QLL melting less frequently. This is expected as the system preferentially explores around basins in the free energy, which in this case is located at the equilibrium QLL thickness. However, even at large $n_{l_{n_i}}$ values the fluctuations in $\Delta F(\mathrm{t})$ are small and reach a steady-state value, indicating that our simulations are well converged.

The convergence of the TIP4P/Ice metadynamics simulations are plotted in Figure 5.11 for the basal and prism surfaces at 260 K. The fluctuations are larger compared to mW indicating that the TIP4P/Ice free energies are less well converged. For large values of $n_{l_{n_i}}$, the TIP4P/Ice $\Delta F(t)$ fluctuations have not yet reached a steady equilibrium value, however, for $n_{l_{n_i}} = 500$ the TIP4P/Ice free energy profiles are well converged. The more time the system spends in a given configuration, the lower its free energy computed from metadynamics. $\Delta F(\mathrm{t})$ is negative at first because the system initially explores large n_l values.

5.4 Chemical potential

The chemical potential difference between a QLL molecule and ice molecule, $\Delta\mu$ is defined as the derivative of the free energy with respect to n_l

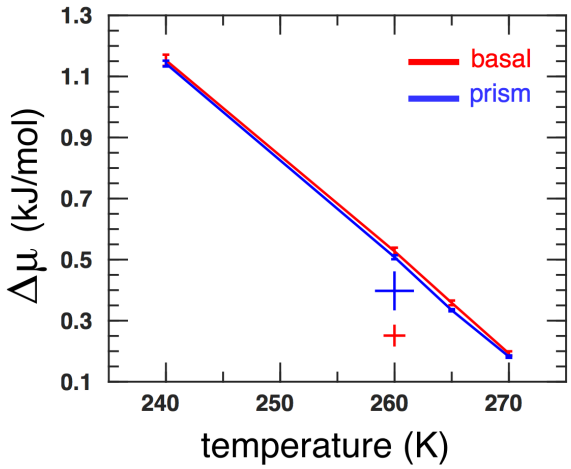


Figure 5.12: Chemical potential difference, $\Delta\mu$ between the QLL and ice calculated from metadynamics simulations at 240 K, 260 K, 265 K and 270 K using mW. The TIP4P/Ice $\Delta\mu$ at 260 K is indicated by a cross with a height corresponding to the standard deviation of the mean.

$$\Delta\mu = \left(\frac{dF}{dn_l}\right)_{T,V,N} \tag{5.2}$$

$\Delta\mu$ is calculated from our metadynamics simulations by fitting a straight line to the free energy profiles. The range of values over which we compute the gradient is determined by visual inspection of the convergence results. We choose the largest range of n_l such that the upper boundary is converged to an equilibrium value and no longer decreasing or increasing. For mW the range used is $n_l = 300$ to 450 at 240 K and $n_l = 350$ to 600 at 260 K, 265 K and 270 K. For TIP4P/Ice at 260 K the range of n_l used to compute $\Delta\mu$ is from 350 to 500.

In the previous sections we described how the free energy is calculated by mediating over 100 free energy reconstructions for mW and 500 for TIP4P/Ice. $\Delta\mu$ is calculated for each free energy reconstruction resulting is 100 estimates for mW at each temperature and 500 estimates for TIP4P/Ice at 260 K. The mean value of $\Delta\mu$ is plotted against temperature in Figure 5.12 for our mW simulation. The error bars indicate the standard deviation of the means. A cross indicates the TIP4P/Ice chemical potential at 260 K and its size corresponds to the standard deviation.

$\Delta\mu$ in Figure 5.12 decreases linearly with temperature which is consistent with both theoretical and experimental results [121, 122]. The theoretical $\Delta\mu$ for growth from liquid is $\Delta\mu \approx \Delta H_{fus}\frac{T_m - T}{T_m}$ where T_m is the melting temperature and ΔH_{fus}

is the enthalpy of fusion [121]. At 260 K, this yields a value $\sim$ 0.29 kJ/mol. At melting point, $\Delta\mu = 0$ which our results extrapolate to. Estimates of $\Delta\mu$ obtained by thermodynamic integration of small I_h clusters using TIP4P/Ice report $\Delta\mu \approx 0.611$ kJ/mol [39] and $\Delta\mu \approx 0.615$ kJ/mol [123] at 230 K which is smaller than our projections. At 260 K, we estimate $\Delta\mu = 0.527$ kJ/mol and 0.508 kJ/mol for the basal and prism mW surfaces, respectively, which is larger than the theoretical value of $\sim$ 0.29 kJ/mol. With TIP4P/Ice we estimate 0.250 kJ/mol and 0.398 kJ/mol for the basal and prism surfaces, respectively. This represents a good agreement with the the theoretical estimate although there are large errors associated with the TIP4P/Ice estimates. Our metadynamics approach overestimates $\Delta\mu$ compared to results from thermodynamic integration and theory [39, 121, 123]. The theoretical approach uses the enthalpy of fusion between water and ice to compute $\Delta\mu$. The structure of ice penetrates into the QLL and quasi-liquid molecules have a degree of crystallinity absent in water. QLL molecules have ice-like properties and therefore the enthalpy of fusion of ice and quasi-liquid molecules can reasonably be assumed to be smaller than ice and water. This suggests that the theoretical approach overestimates $\Delta\mu$. Based on these considerations, the TIP4P/Ice $\Delta\mu$ offers a better agreement with the experimental value than the mW estimates.

Our metadynamics results reveal that the chemical potential driving force for melting the mW QLL at equilibrium is greater than for TIP4P/Ice. In Chapter 4 we demonstrate that the mW QLL is over-structured and forms a small QLL. The mW QLL size fluctuates with smaller amplitudes compared to the TIP4P/Ice QLL. These observations correlate with our metadynamics results which show that the energy cost of melting the mW QLL is greater than the TIP4P/Ice QLL. Whilst $\Delta\mu$ estimated for the basal and prism mW surfaces are comparable at each temperature, the TIP4P/Ice estimates are different. There are large uncertainties associated with the TIP4P/Ice estimates and it is unclear if longer simulations would converge the basal and prism values.

5.5 Interface Energy

The metadynamics free energies can be explained from a surface free energy viewpoint if we consider that the ice/vapour interface can be split into two interfaces: the ice/QLL and the QLL/vapour interfaces [35]. The surface free energy, γ can therefore be expressed in terms of the ice/QLL interface free energy, γ_{SL}, the QLL/vapour interface free energy, γ_{LV} and the ice/vapour interface free energy, γ_{SV} as

$$\begin{aligned}\gamma(d) &= \gamma_{SV} + f(d)(\gamma_{SL} + \gamma_{LV} - \gamma_{SV}) \\ &= \gamma_{SV} + f(d)\Delta\gamma\end{aligned} \tag{5.3}$$

where $f(d)$ is a function dependent on the QLL thickness d. γ_{SV} is the theoretical interface free energy when there is no surface melting and cannot be measured experimentally whenever there is a QLL on the surface of ice. $f(d) \to 0$ as $d \to 0$ and $f(d) \to 1$ as $d \to \infty$. As the QLL thickness vanishes, the surface free energy is equal to the ice/vapour interface free energy, $\gamma(d = 0) = \gamma_{SV}$. In the limiting case, as $d \to \infty$, the surface free energy is the sum of the pure ice-liquid and pure liquid-vapour interfaces, $\gamma(d \to \infty) = \gamma_{SL} + \gamma_{LV}$. Surface melting is thermodynamically favourable when $\Delta\gamma < 0$.

Following from this, the free energy per unit area as a function of d can be expressed as

$$F(d) = \frac{\Delta H_f^m \rho}{T_m} d\Delta T + \gamma(d) \tag{5.4}$$

where ΔH_f^m is the enthalpy of fusion per unit mass, ρ is the density of water and $\Delta T = T_m - T$ is the temperature below the melting temperature, T_m. The first term quantifies the free energy cost of melting the QLL below T_m. The second competing term quantifies the stabilising effect of surface melting. The functional form of Eq. 5.4 describes our metadynamics free energies in Figures 5.2 and 5.5 if we consider that d is directly correlated with n_l. When d is large enough, $f(d)$ is close to 1 and

the free energy increases linearly with d. In this region, the gradient of the free energy decreases with temperature, as observed in our calculations of $\Delta\mu$.

The functional form of $f(d)$ depends on the surface force acting between the solid-liquid and liquid-vapour interfaces. Modelling the forces using a Van der Waals functional form, as suggested in Ref. [124], yields $f(d) = \frac{d^2}{(d^2+\sigma^2)}$. In general, power law interaction potentials can be modelled using $f(d) = \frac{d^n}{(d^n+\sigma^n)}$, where n depends on the nature of the interactions and σ is of the order of molecular diameter [125]. If the surface force acting between the two interfaces is exponentially decaying, $f(d) = 1 - \exp(-d/\lambda)$.

Differentiating $F(d)$ allows us to estimate $\Delta\mu$ and quantify the agreement of our simulations with theory:

$$\begin{aligned} F'(d) &= \frac{\Delta H_f^m \rho}{T_m}\Delta T + f'\Delta\gamma \\ &= \frac{\Delta H_f^m \rho}{T_m}\Delta T - \frac{1}{\lambda}\exp\left(\frac{-d}{\lambda}\right)\Delta\gamma. \end{aligned} \tag{5.5}$$

When $d \gg \lambda$, the second term vanishes and $\Delta\mu\rho = \Delta H_f^m \rho \Delta T/T_m$ resulting in $\Delta\mu \propto \Delta T$, as observed in our metadynamics calculations.

The equilibrium QLL thickness occurs at the free energy minimum and therefore can be estimated by equating Eq. 5.5 to 0. This leads to a logarithmic relationship between the equilibrium QLL thickness and temperature

$$d(T) = \lambda \ln\left(\frac{\Delta\gamma T_m}{\Delta H_f^m \rho\lambda\Delta T}\right) \tag{5.6}$$

where λ is the intermolecular interactions decay length which can be measured experimentally. This can be expressed as

$$\begin{aligned} d(T) &= \lambda \ln\left(\frac{\Delta\gamma T_m}{\Delta H_f^m \rho\lambda}\right) - \lambda \ln \Delta T \\ &= a - b \ln \Delta T \end{aligned} \tag{5.7}$$

which agrees with the experimental functional form in Eq. 1.4 and 1.5, our TIP4P/Ice simulation result in Figure 4.11, and previous experimental results [1, 33–35].

This model of the energetics of the surface of ice assumes that the QLL structure is similar to bulk water. If this is not the case, γ_{SL} and γ_{LV} should be considered as parameters of the model rather than thermodynamic quantities. If the QLL structure is not homogeneous and gradually changes with depth, a distinct ice/QLL interface cannot be defined and the model does not stand. However, the functional form of Eq. 5.7 agrees with both experimental findings and our simulation results suggesting the assumptions of the model are reasonable.

5.6 Conclusions

Our results in Chapter 4 demonstrate that mW is poor at capturing the kinetics and structure of the QLL. The lack of hydrogen atoms results in fast molecular mobility and large diffusion coefficients. However, our metadynamics simulations show that mW is a good water model for calculating thermodynamic quantities. Studies using mW have shown that mW reproduces energetics and thermodynamic anomalies well [94].

The free energies recovered using metadynamics as a function of n_l show that a QLL will form on the surface of ice with an equilibrium thickness. Melting the QLL beyond its equilibrium is unfavourable but recrystallising it if it is larger than equilibrium is favourable, and there are no metastable states. These results have important implications for a vapour molecule condensing on the QLL surface. If the QLL is larger than its equilibrium, we can infer that crystallisation will occur spontaneously and the adsorption of vapour will result in the crystallisation of QLL molecules.

$\Delta\mu$ provides a measure of the driving force of crystallisation and melting. Our results show that this driving force is equivalent for both crystallographic surfaces using mW. The energy required to melt an ice molecule is influenced by the hydrogen bonding network. This network is identical for a prism and basal bulk ice molecule.

The hydrogen bonding networks of the mW and TIP4P/Ice QLLs are different and therefore the two models yield different free energy profiles and $\Delta\mu$ estimates. In order to understand the dynamics of QLL crystallisation and the impact of $\Delta\mu$, we perform unbiased MD simulations of QLL crystallisation.

Chapter 6

QLL Crystallisation

The metadynamics results in Chapter 5 indicate that QLL crystallisation occurs spontaneously when the QLL is larger than its equilibrium size. In this Chapter we investigate this further using molecular dynamics simulations. The crystallisation of the QLL can be simulated using MD by first artificially melting a surface so that the QLL size is much larger than at equilibrium. MD simulations are then performed, and the evolution of the number of QLL molecules on the melted surface is monitored as the QLL crystallises.

6.1 TIP4P/Ice Simulations

An artificially melted TIP4P/Ice surface is prepared using MD by simulating molecules on one surface for 2 ns at 300 K in the NVT ensemble whilst holding the remaining atoms frozen in their positions. Details of the thermostatting and general MD setup used are outlined in Section 3.3. A surface region corresponding to 8 molecular bilayers and a thickness ~ 35 Å is melted during this procedure. Once melted, the supercooled QLL is simulated for 1 ns to decrease its temperature back down to its initial value. Finally, the QLL/ice interface is allowed to relax by running MD on all atoms for a total of 2M steps using a time step increasing from 0.001 fs to 1 fs. During the 2M steps, ice molecules at the interface with the QLL are able to readjust their positions, however, the small time step ensures that the system does

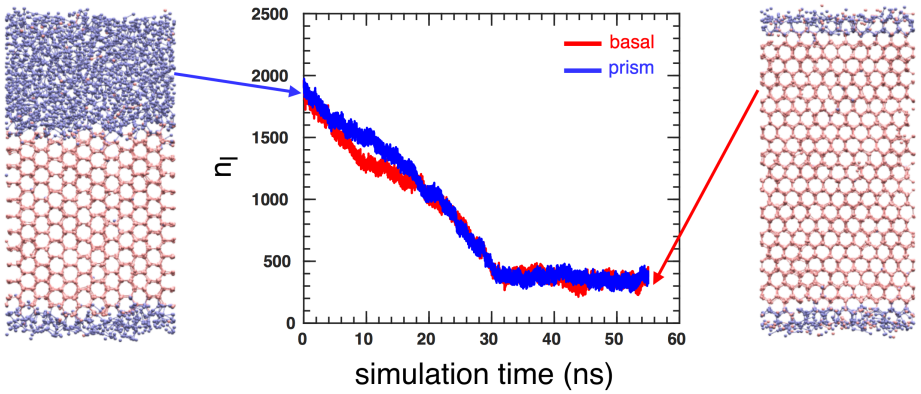


Figure 6.1: Evolution of the number of QLL molecules during the TIP4P/Ice crystallisation simulations at 260 K. Snapshot on the left shows the initial prism configuration whilst the snapshot on the right shows the final basal configuration, following 55 ns of simulations.

not recrystallise. NVT simulations are then performed for up to 160 ns until the QLL has recrystallised and returned back to equilibrium.

Figure 6.1 shows the evolution of n_l for the basal and prism crystallising QLLs at 260 K, using TIP4P/Ice. The initial prism configuration is shown in the snapshot on the left. Clearly the artificially melted QLL on the top surface is much larger than the equilibrium QLL on the bottom surface. The snapshot on the right is the final basal configuration at the end of the simulation. The top surface which was initially melted artificially is now indistinguishable from the bottom surface, indicating that crystallisation has occurred and returned the melted QLL back to equilibrium. In both snapshots QLL molecules are coloured pink and ice molecules are coloured blue, based on the q_3 parameter. n_l is plotted against time in the central plot in Figure 6.1. The results show that QLL crystallisation occurs at a constant rate and returns the QLL back to equilibrium. Growth is continuous and the evolution of n_l is linear. There are no metastable states which is consistent with our metadynamics simulations which showed no free energy minima other than at the equilibrium QLL size.

In order to understand the effect of temperature on QLL crystallisation, TIP4P/Ice QLL crystallisation simulations of the basal and prism surfaces are performed at 240 K, 245 K, 250 K, 255 K, 260 K, 265 K and 270 K. The simulations are repeated three times at each temperature to ensure a reliable and reproducible result. This is

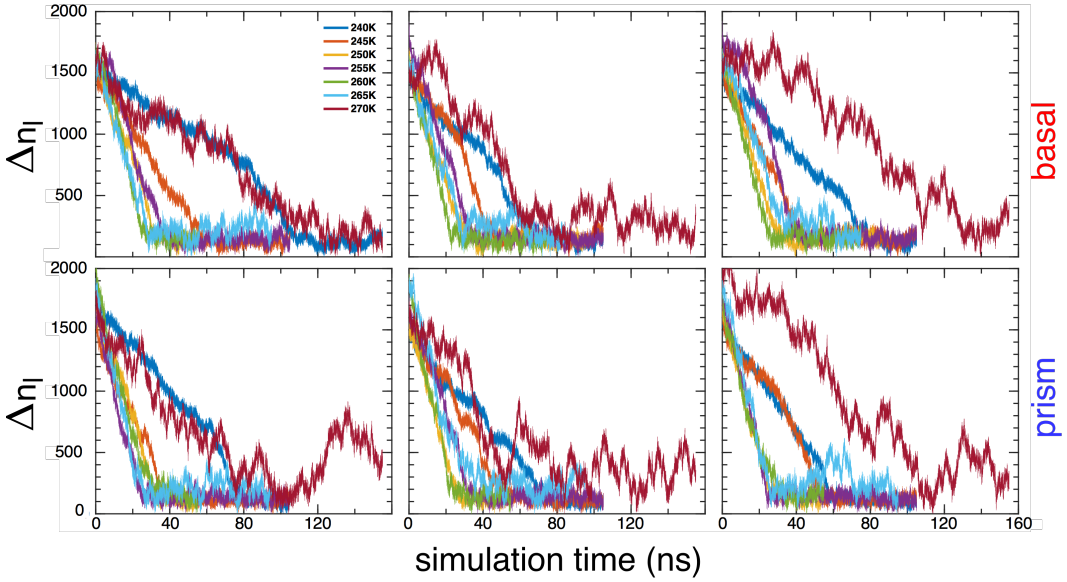


Figure 6.2: Evolution of Δn_l during the TIP4P/Ice QLL crystallisation simulations of the basal surface (top row) and prism surface (bottom row) for the three repeats.

achieved by initialising the velocities in both the QLL melting and QLL/ice interface relaxation steps with different random number seeds. Δn_l is calculated from the simulations using $\Delta n_l = n_l - \min(n_l)$ and is plotted for both surfaces and for the three repeats in Figure 6.2. The plots of Δn_l show that the rate of QLL crystallisation depends on temperature.

The rate of QLL crystallisation can be estimated from the slope of Δn_l using linear regression. In Figure 6.3 Δn_l is plotted for one of the basal and prism repeats at each of the simulated temperatures. The plots include a linear fit in red which can be used to estimate the rate of growth. The linear fits are a reasonable approximation at all temperatures, although at 270 K Δn_l fluctuates during crystallisation.

The linear growth rate of TIP4P/Ice QLL crystallisation is estimated from linear fits to Δn_l and plotted against temperature in Figure 6.4. The lengths of the slabs are taken into account to express the the linear growth rates in units of Å/ns. The slab length is computed from the mean coordinate of the outermost 100 molecules on each surface during MD simulations. Averaging over 100 molecules takes into account half a bilayer and ensures that our estimate of the slab length is not biased by evaporating molecules. The rate is averaged over the three repeats and the error bars in the plots indicate the standard deviation. The results show that as temperature is

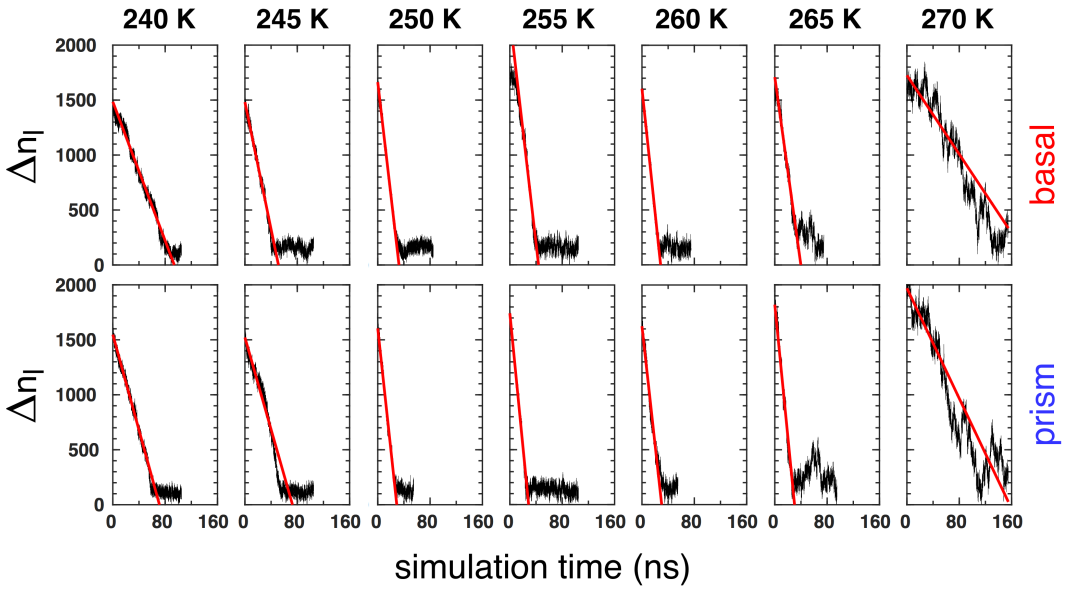


Figure 6.3: Evolution of Δn_l during the TIP4P/Ice QLL crystallisation simulations of the basal surface (top row) and prism surface (bottom row) at a 240 K, 245 K, 250 K, 255 K, 260 K, 265 K and 270 K. Red lines are linear fits to the slopes.

increased from 240 K to 270 K, the rate increases then decreases. At 240 K and 270 K, the rates of growth are slow and quantitatively similar. The basal and prism QLL crystallisation rates are subject to a similar temperature dependence and overlap within the standard deviation. Within the margin of error, it is not clear whether the maxima in the basal and prism rates occur at different temperatures.

At 250 K, we estimate linear growth rates of 0.73 ± 0.13 Å/ns and 0.83 ± 0.099 Å/ns for the basal and prism surfaces, respectively. At 260 K, we estimate growth rates of 0.98 ± 0.090 Å/ns and 0.92 ± 0.069 Å/ns for the basal and prism surfaces,

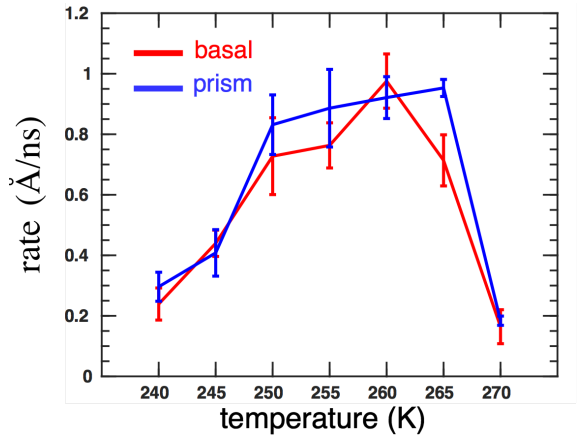


Figure 6.4: Rate of QLL crystallisation of the basal (red) and prism (blue) TIP4P/Ice surfaces.

respectively. This is in good agreement with the recent TIP4P/Ice simulations using the seeding technique which estimate 0.633 Å/ns and at 0.933 Å/ns for the secondary prism surface at 250 K and 260 K, respectively [39]. It is also in reasonable agreement with non surface specific experimental measurements of the linear growth rate of ice in supercooled water which estimate close to 0.6 Å/ns at 260 K [126, 127]. The simulations paint a complex picture of QLL crystallisation. In order to understand the effects of hydrogen atoms and water model on the observed rate trend, QLL crystallisation simulations are performed using mW.

6.2 mW Simulations

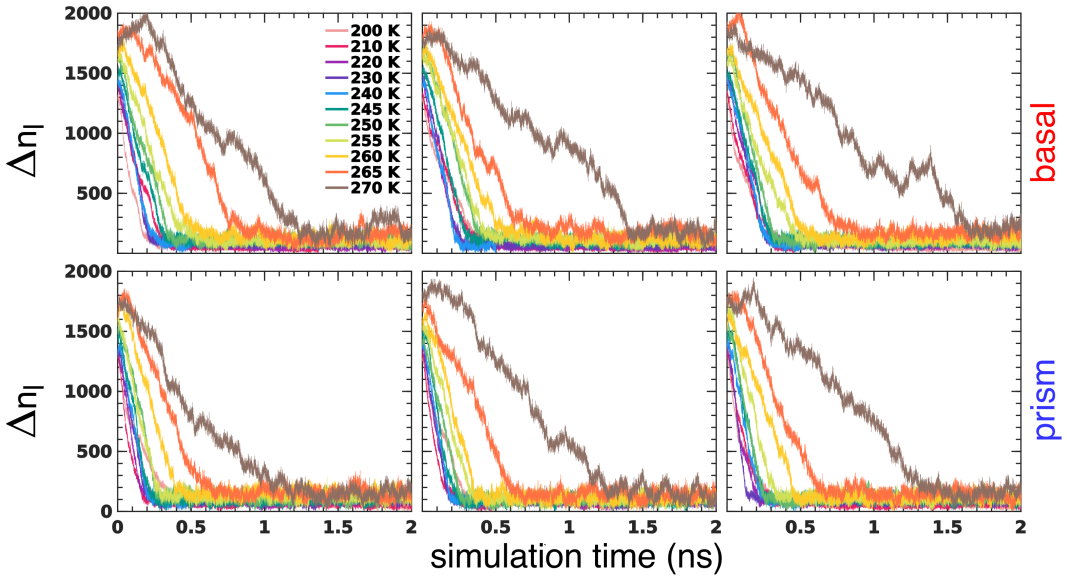


Figure 6.5: Evolution of n_l during the mW QLL crystallisation simulations of the basal surface (top row) and prism surface (bottom row) for three of the ten repeats.

QLL crystallisation simulations are performed with mW by first preparing artificially melted surfaces using the same procedure described for TIP4P/Ice. MD simulations are then run for up to 3 ns at 200 K, 210 K, 220 K, 230 K, 240 K, 245 K, 250 K, 255 K, 260 K, 265 K and 270 K. Shorter simulations are performed because the QLL crystallisations occurs much more rapidly than with TIP4P/Ice. The simulations are repeated 10 times at each temperature by changing the random number seeds for velocity initialisation. Details of the MD setup and thermostat are

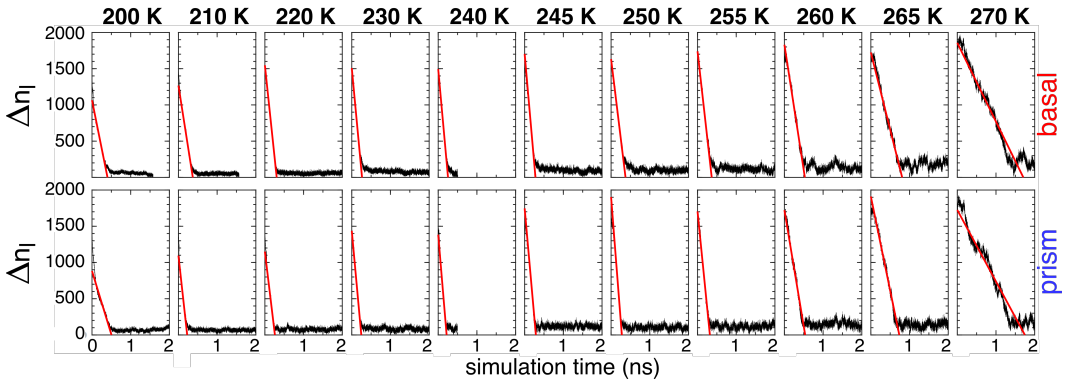


Figure 6.6: Evolution of n_l during the mW QLL crystallisation simulations of the basal surface (top row) and prism surface (bottom row) at a 200 K, 210 K, 220 K, 230 K, 240 K, 245 K, 250 K, 255 K, 260 K, 265 K and 270 K. Red lines are linear fits to the slopes.

outlined in Section 3.3. Figure 6.5 is a plot of Δn_l for three of the ten repeats at each of the simulated temperatures. The evolution of the mW QLL reaffirms the TIP4P/Ice result: constant QLL crystallisation occurs with no metastable states. The rate of crystallisation depends on temperature as observed with TIP4P/Ice, however, the timescale of the mW QLL crystallisation is significantly faster than with TIP4P/Ice. A consistent result is obtained across the ten repeats at each temperature and for both surfaces.

The mW rate of crystallisation is estimated from linear fits to the slope of Δn_l. Figure 6.6 shows the fits to Δn_l in red for one the mW repeats at each of the simulated temperatures. It is clear from the plots that a linear fit can accurately describe the evolution of the number of QLL molecules during the mW crystallisation process.

The rate of mW QLL crystallisation, estimated from linear fits to Δn_l, is plotted in Figure 6.7 against temperature. The lengths of the slabs are taken into account to express the linear growth rates in Å/ns. The temperature dependence of the rate obeys a similar functional form to that of TIP4P/Ice, however, it spans a much larger range of temperatures. Whilst the TIP4P/Ice basal and prism QLL crystallisation rates peak at $\sim$ 260 K, the mW rates peak at $\sim$ 230 K. Recently published simulation work using the seeding technique with spherical I_h clusters reports a maximum growth rate at 260 K and 230 K for the secondary prism surface using TIP4P/Ice and mW, respectively [39]. This is in direct agreement with our results strongly indicating

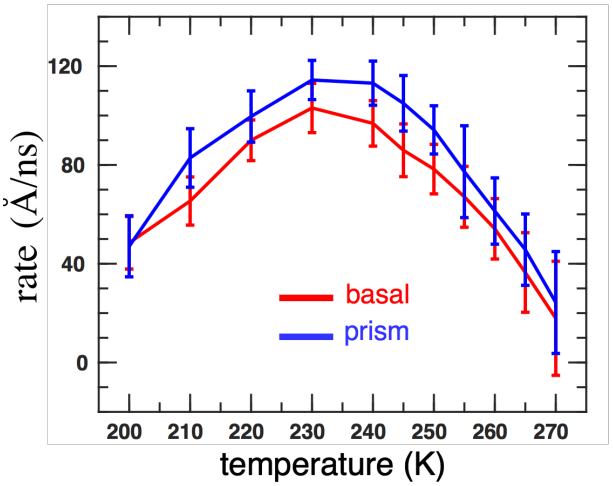


Figure 6.7: Rate of mW QLL crystallisation of the basal (red) and prism (blue) surfaces

that the temperature of maximum rate of growth is independent of crystallographic surface but dependent on the water model used. MD simulations of the crystallisation of supercooled water at ice interfaces in the absence of vacuum report a decrease in temperature of the maximum growth rate by 1 K as the time step of MD simulations is increased from 1 fs to 2 fs and a decrease by 2 K as the time step is increased from 1 fs to 3 fs [40]. The time step is found to affect the free energy of the bulk phase and this demonstrates that the maximum growth rate temperature is sensitive to the parameters used.

Simulations of supercooled water crystallisation by Rozmanov [40] are performed for up to 25 ns using TIP4P/2005. The rate of crystallisation is estimated by identifying the interface position from the Fourier transform of the density profile. The findings show that the rate of crystallisation of the prism ice surface is 30% faster than the basal surface. This is consistent with our mW rates of crystallisation. Using TIP4P/Ice, our basal and prism rates are numerically much closer to one another, although the prism surface appears to crystallise faster than the basal surface at a number of temperatures. The MD simulations described in Chapter 4 show that the equilibrium prism QLL is larger than the equilibrium basal QLL with mW but that the two QLLs are equivalent using TIP4P/Ice. This suggests a complex relationship between the equilibrium QLL and the rate of QLL crystallisation. The mW rate of QLL crystallisation is two orders of magnitude faster than TIP4P/Ice. This is

in general agreement with the work of Espinosa [39] which reports an mW rate of growth three orders of magnitude faster than TIP4P/Ice for the secondary prism surface. The lack of hydrogen atoms in the mW model results in lower diffusion energy barriers and faster dynamics [94]. We observe much faster dynamics using mW, as shown by our self-diffusion coefficient measurements, and significantly larger rates of QLL crystallisation compared to TIP4P/Ice.

Crystallisation of the QLL results in the release of latent heat at the interface which hinders crystallisation and must be dissipated to enable crystallisation to continue. The mechanism through which heat is removed from the interface during crystallisation is complex. The simulation setup used in this work couples every molecule to the thermostat heat bath resulting in an isothermal system and therefore care must be taken when comparing the results with experiments. Nonetheless, the temperature dependence of the rates of QLL crystallisation suggest an interplay between two competing effects. One which acts to increase the rate with temperature and one which decreases it with temperature. In the following section we outline a theoretical description of the rate of QLL crystallisation and explain the temperature dependence.

6.3 Discussion

The dynamics of ice crystal growth depend on an interplay of attachment kinetics, particle diffusion and heat diffusion. Particle diffusion transports water molecules to the growing crystal, whereas heat diffusion removes the latent heat generated by solidification [11]. Particle transport is defined by

$$\frac{\mathrm{d}c}{\mathrm{d}t} = D\nabla^2 c, \tag{6.1}$$

where c is the number density of water molecules surrounding the crystal and D is the diffusion constant. Crystal growth of an ice surface in contact with supercooled water is driven by the chemical potential difference at the interface.

A liquid molecule in contact with the solid must change orientation to join the

crystalline order. Surrounding molecules, however, hinder the free motion of the molecule and limit it to vibrations of frequency ν, of the order of lattice vibrations. In order to crystallise, the liquid molecule must overcome the energy barrier of molecular diffusion, E_d. At a given temperature, T a molecule acquires the energy to overcome the diffusion barrier with a probability proportional to the Boltzmann weight $\exp(-E_d/k_BT)$. The rate of crystallisation of a liquid water molecule in contact with a crystalline ice surface is given by $\nu \exp(-E_d/k_BT)$. However, the process is an equilibrium which also involves the melting of crystal molecules. Below the melting temperature the chemical potential of the liquid is greater than the solid by $\Delta\mu$, and the rate of crystallisation is greater than that of melting by a factor $\exp(-\Delta\mu/k_BT)$. The solidification front increases by a molecular diameter a when a molecule crystallises, allowing the growth rate to be expressed as

$$v_n = a\nu \exp\left(-\frac{E_d}{k_BT}\right)\left[1 - \exp\left(-\frac{\Delta\mu}{k_BT}\right)\right] \tag{6.2}$$

The diffusion constant D can be expressed in terms of the viscosity η using the Einstein-Stokes relation

$$D = \frac{k_BT}{6\pi\eta a} = \nu a^2 \exp\left(-\frac{E_d}{k_BT}\right) \tag{6.3}$$

leading to

$$v_n = \frac{D}{a}\left[1 - \exp\left(-\frac{\Delta\mu}{k_BT}\right)\right] \tag{6.4}$$

which is known as the Wilson-Frenkel formula [121]. D provides a measure of the speed of diffusion of liquid molecules into sites where they can join the crystal lattice. The Wilson-Frenkel growth velocity in Eq. 6.4 relates the rate of crystal growth to the product of molecular diffusion and chemical potential.

Molecular diffusion increases with temperature, as observed in our calculations of the translational and rotational diffusivities plotted in Figures 4.31 and 4.34, respectively. The mW translational diffusion of the prism surface is greater than the basal surface, therefore, the Wilson-Frenkel equation predicts a faster prism rate of

QLL crystallisation, as observed in our simulations. The TIP4P/Ice basal and prism diffusion coefficients are equivalent and overlap within the standard deviation which explains the numerical proximity of their rates of QLL crystallisation.

Inspired by the Wilson-Frenkel equation, we fit our TIP4P/Ice QLL rate of crystallisation to the equation

$$r = \frac{D}{\delta_l}\left[1 - \exp\left(\frac{-\lambda\Delta\mu}{RT}\right)\right] \tag{6.5}$$

where D is the TIP4P/Ice translational diffusion coefficient, calculated in Section 4.5.1 from MD simulations, $\Delta\mu$ is the difference in chemical potential calculated of a QLL and ice mW molecule, calculated from mW metadynamics simulations in Section 5.4, and λ is a scaling parameter which takes into account that the TIP4P/Ice $\Delta\mu$ is smaller than the mW $\Delta\mu$. δ_l is the QLL thickness obtained from the TIP4P/Ice QLL size, $\bar{n}_l$ as $\delta_l = l_z\bar{n}_l/n_{tot}$, where l_z is the length of the slab and n_{tot} is the total number of molecules in the slab. In this construct l_z/n_{tot} is a measure of the molecular diameter, a. D and δ_l are calculated from TIP4P/Ice MD simulations whereas $\Delta\mu$ is calculated from mW metadynamics simulations. λ is the only parameter of the model.

$\Delta\mu$ is computed from mW metadynamics simulations at 240 K, 260 K, 265 K and 270 K for the basal and prism surfaces in Section 5.4. The trend shows a linear dependence on temperature and a very weak dependence on the crystallographic surface. The values of $\Delta\mu$ at 245 K, 250 K and 255 K are interpolated from linear fits to the basal and prism mW $\Delta\mu$. The interpolated basal and prism $\Delta\mu$ results are averaged to give a non surface specific estimate that is used in Eq. 6.5 to compute the rate of basal and prism QLL crystallisation.

The rate of TIP4P/Ice QLL crystallisation is fitted to Eq. 6.5 and plotted in Figure 6.8. The results show that Eq. 6.5 describes the rate of QLL crystallisation well. The scaling parameters obtained from the fits are $\lambda = 0.40 \pm 0.12$ and 0.44 ± 0.13 for the basal and prism surfaces, respectively. The uncertainty indicates the 95% confidence interval. $\lambda < 1$ indicates that the TIP4P/Ice $\Delta\mu$ is less than the mW $\Delta\mu$ which is consistent with our metadynamics results at 260 K. At 260 K, the

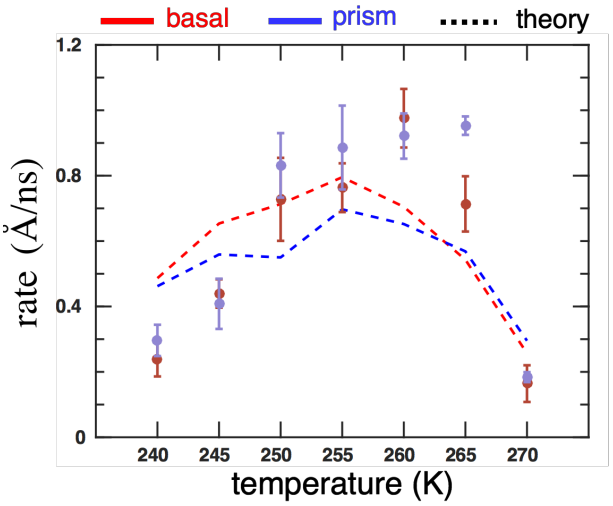


Figure 6.8: Rate of TIP4P/Ice QLL crystallisation estimated from our modification of the Wilson-Frenkel equation.

ratio of the TIP4P/Ice to mW $\Delta\mu$, based on our metadynamics results in Section 5.4, yields 0.47 and 0.78 for the basal and prism surfaces, respectively. This is in reasonable agreement with the fitting parameter from our model of the rate of QLL crystallisation.

6.4 Saturation Vapour Pressure: An interplay of $\Delta\mu$, QLL Structure and Diffusivity?

An ice or water surface in air is in equilibrium with the vapour phase. At equilibrium the vapour pressure is known as the saturation vapour pressure, and the number of molecules joining and leaving the vapour phase are equal. The larger the number of vapour molecules, the larger the vapour pressure. As temperature increases, the saturation vapour pressure increases because vibrating ice or water molecules are able to escape more easily. Ice molecules are more strongly bonded than water molecules and therefore can escape less easily. As a result, the saturation vapour pressure of supercooled water is greater than ice at a given temperature.

In general, a metastable phase characterised by a high Gibbs energy will have a higher vapour pressure than a stable phase with a low Gibbs energy. Below the melting point, supercooled water has a higher Gibbs energy than ice and therefore a

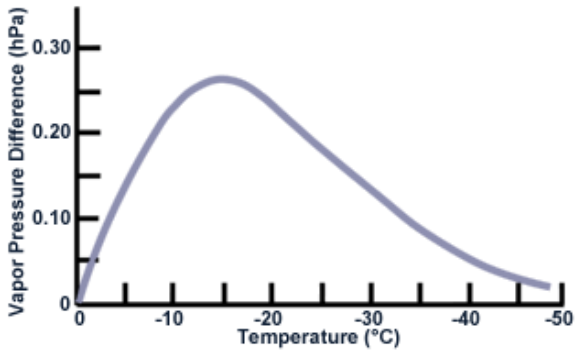


Figure 6.9: Difference in saturation vapour pressure of supercooled water and ice [128].

larger saturation vapour pressure. The difference in the saturation vapour pressure of supercooled water ρ_w and ice ρ_i is defined as $\Delta\rho = \rho_w - \rho_i$. $\Delta\rho$ is plotted in Figure 6.9 at temperatures where a QLL forms on the surface of ice. $\Delta\rho$ is always positive as the saturation vapour pressure of supercooled water is greater than ice. However, the shape of the plot with temperature remains unexplained.

Vapour pressure can be described in relation to the chemical potential of the vapour phase, μ_v. Assuming the vapour phase behaves as an ideal gas, the dependence of μ_v on pressure p can be expressed as

$$\mu_v = \mu_v^\circ + k_B T \ln\left(\frac{p}{p^\circ}\right) \tag{6.6}$$

where μ_v° is the chemical potential at standard state pressure, p°. At saturation vapour pressure, μ_v is equal to the chemical potential of the second phase. For supercooled water and ice, this leads to

$$\begin{aligned} \mu_v^\circ + k_B T \ln\left(\frac{p_l}{p^\circ}\right) &= \mu_l \\ \mu_v^\circ + k_B T \ln\left(\frac{p_s}{p^\circ}\right) &= \mu_s \end{aligned} \tag{6.7}$$

where μ_l and μ_s are the chemical potentials of supercooled water and ice, respectively, and p_l and p_s are the equilibrium vapour pressures of supercooled water and ice, respectively. The difference of the equalities in Eq. 6.7 leads to

$$k_B T \ln\left(\frac{p_l}{p_s}\right) = \Delta\mu_{sl} \tag{6.8}$$

where $\Delta\mu_{sl} = \mu_l - \mu_s$. Below the melting point, $\Delta\mu_{sl} > 0$ and $p_l > p_s$. At the melting point $\Delta\mu_{sl} = 0$ leading to $p_l = p_s$ and $\Delta\rho = 0$. Chemical potential considerations explain the observed decrease in $\Delta\rho$ as the melting is approached, shown in Figure 6.9. This result can equally be rationalised by a QLL size which diverges and becomes more and more liquid-like close to the melting point. Eq. 6.8 can be rearranged as

$$p_l = \exp\left(\frac{\Delta\mu_{sl}}{k_B T}\right) p_s \tag{6.9}$$

$\Delta\mu_{sl}$ decreases with temperature, therefore, Eq. 6.9 indicates that $\Delta\rho$ increases with temperature. However, at 240 K Figure 6.9 shows that $\Delta\rho$ is small and heading to zero with decreasing temperature. The QLL size decreases with temperature, therefore, the low temperature $\Delta\rho$ result cannot be rationalised in terms of a liquid-like QLL. Interestingly, we notice that the trend in $\Delta\rho$ is similar to our QLL crystallisation rate from TIP4P/Ice simulations. A second effect competing with $\Delta\mu_{sl}$ appears to result in a maximum in $\Delta\rho$ at a similar temperature to the maximum in the QLL rate of crystallisation.

In the previous Section we describe through the Wilson-Frenkel formula that the crystallisation of a QLL molecule involves its transportation to sites where it may crystallise. Decreased diffusion hinders the process, and a similar effect may impact the desorption of a QLL molecule. At low temperatures QLL crystallisation is slow which could impact the stability of the surface and increase its Gibbs energy. Previous simulation and experimental works suggest that at low temperatures the QLL is formed of one bilayer whilst at higher temperatures it is composed of two or more bilayers [7, 36–38, 66]. A recent simulation and experimental work suggests that the transition from one melted bilayer to two melted bilayers occurs at 257 K [66]. This temperature coincides with the maximum of $\Delta\rho$. An intrinsic link between QLL structure, diffusivity and vapour pressure remains speculative but warrants further investigation by the community.

6.5 Conclusions

The kinetics of QLL crystallisation is governed by the atomic structure and dynamics. Previous theoretical and simulation works find that the key factors are the diffusivity and size of the QLL, and the chemical potential change as the surface adds bilayers [12, 129, 130]. Our results affirm these observations and show that the rate of QLL crystallisation depends on an interplay of the thermodynamic driving force, described by $\Delta\mu$, and the molecular diffusivity. $\Delta\mu$ decreases with temperature whilst the diffusivity increases with temperature. Previous simulation studies of the growth of ice fit the rate of crystallisation to arbitrary functions with little or no theoretical justification [39, 40]. We demonstrate that the QLL crystallisation can be successfully described using the diffusion and chemical potential computed from MD and metadynamics simulations. Our results validate the Wilson-Frenkel description of ice growth implying that crystallisation involves the diffusion of QLL molecules into sites where they may crystallise. Our TIP4P/Ice QLL rate estimates correlate with previous simulation and experimental results [39, 126, 127], whereas the mW rate of QLL crystallisation is two orders of magnitude faster. The lack of hydrogens in mW results in fast dynamics and rapid QLL crystallisation as molecules are not slowed down by the reorientation of hydrogen atoms.

Using insights from our QLL crystallisation simulations we speculate on a connection between QLL diffusivity, structure and crystallisation rate, and the observed difference in saturation vapour pressure of supercooled water and ice. Vapour pressure gradients at the surfaces of ice crystals and supercooled water droplets govern the diffusion of vapour molecules within clouds and impact crystal growth. A detailed understanding of $\Delta\rho$ could, therefore, have important implications for climate science and warrants further investigation.

Overall our QLL crystallisation simulations provide a description of the kinetics that a crystallising vapour molecule experiences, however, it does not explain the anisotropy of the basal and prism crystallisation rates observed in nature as temperature fluctuates. The QLL/vapour and QLL/ice interfaces are our next focus in our endeavour to explain the basal and prism growth rates in vapour.

Chapter 7

Energetics and Dynamics at the QLL/Vapour & QLL/Ice interfaces

Ice growth in vapour is mediated by a QLL which varies with temperature and has a complex structure and dynamics. Our MD results in Chapter 4 show that the QLL structure is disordered whereas ice is ordered. The QLL thickness increases with temperature but depends weakly on the crystallographic plane using TIP4P/Ice. Analysis of the dangling bonds and molecular orientation in Chapter 4 demonstrate that the outer QLL structure of the basal and prism surfaces are closely related. These similarities suggest that the QLL/vapour interface is not the key factor in determining the relative rate of growth of the two surfaces. In this Chapter we investigate this further by studying the energetics at the QLL/vapour interface using metadynamics. The orientation layer-layer heat maps in Chapter 4 indicate that the basal and prism ice sublayers have distinct structure. At low temperatures when the QLL is small, the in-plane diffusivity of QLL molecules is anisotropic and influenced by the underlying ice structure [38, 47, 48]. The dynamics at the QLL/ice interface are therefore influenced by the ice structure and could be the origin of the anisotropy of the basal and prism growth rates. In the first part of this Chapter we investigate in detail the kinetics at the QLL/ice interface during our TIP4P/Ice MD simulations.

7.1 QLL/Ice interface

Water molecules form four hydrogen bonds within ice bilayers. In the QLL, water molecules form on average three hydrogen bonds in the sublayer at the QLL/ice interface, and two hydrogen bonds in the sublayer at the QLL/vapour interface [29]. MD simulations show that molecules in the ice bilayer at the QLL/ice interface are capable of moving out of their lattice positions, whereas molecules belonging to deeper ice bilayers are not [29]. The number of hydrogen bonds impacts the diffusivity of water, however, there has been no concerted effort to quantify the effects of temperature and ice structure on the kinetics at the QLL/ice interface. In the following, we reveal the impact of temperature on diffusion at the basal and prism QLL/ice interfaces.

The free energy of diffusion normal to the surface, $\Delta F(z)$ can be calculated from the density profiles as

$$\Delta F(z) = -k_B T \log \rho(z) \tag{7.1}$$

where $\rho(z)$ is the density profile along the z-axis normal to the surface, T is the temperature and k_B is the Boltzmann constant. $\Delta F(z)$ is the free energy profile associated with molecular diffusion along the slab and the result is plotted in Figure 7.1. The figure also indicates $\rho(z)$ used to calculate $\Delta F(z)$. It is clear from the plots of $\Delta F(z)$ that a molecule diffusing through the basal and prism surfaces will encounter different energy barriers and will therefore experience different dynamics.

The diffusion barriers impact the dynamics at the QLL/ice interface. In order to study this effect we must first determine the position of the interface in our simulations. $\rho(z)$ and $\Delta F(z)$ allow us to define the QLL/ice interface using two different approaches.

Our first approach defines the QLL/ice interface as the density minimum between the innermost QLL and outermost ice $\rho(z)$ peaks. The innermost QLL $\rho(z)$ peaks are coloured cyan in Figure 7.1 and correspond to the innermost density peaks consisting of more than 75% liquid-like molecules, based on the q_3 parameter. The neighbouring

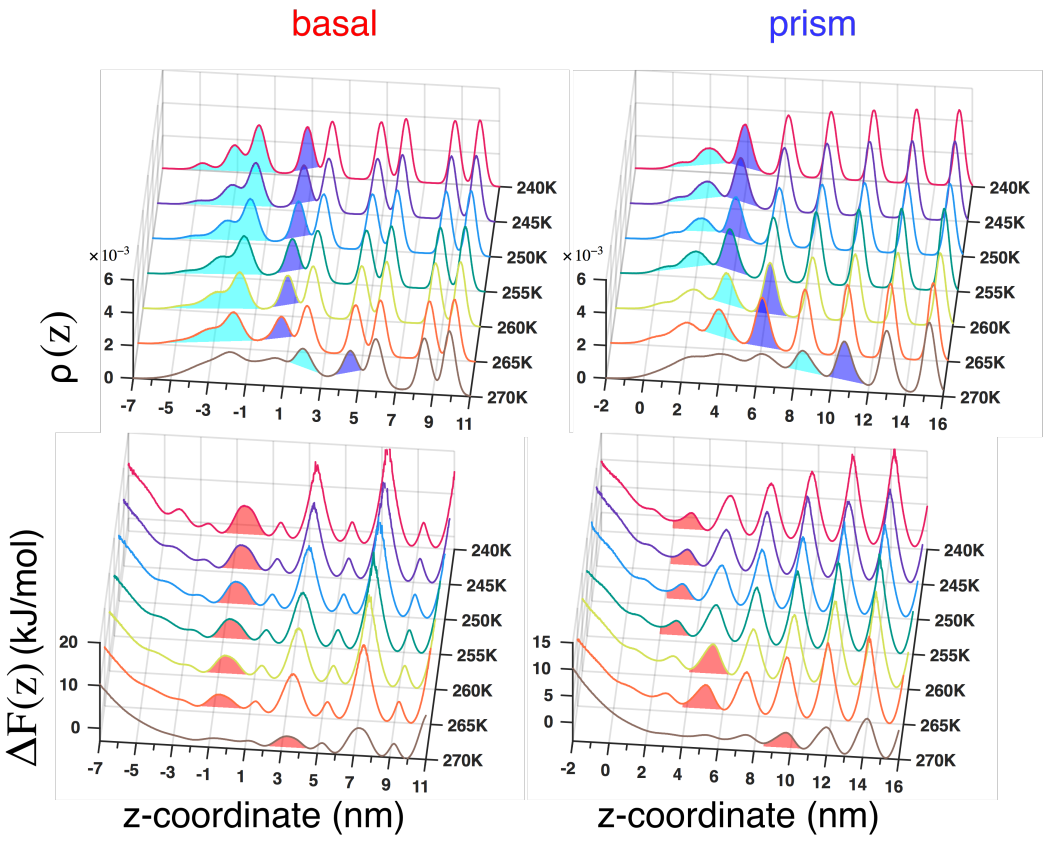


Figure 7.1: Density profiles, $\rho(z)$ and the free energy of diffusion, $\Delta F(z)$ normal to the surface for for the basal and prism TIP4P/Ice surfaces. $\Delta F(z)$ is calculated from the logarithm of $\rho(z)$. Cyan density peaks correspond to the innermost QLL region defined by q_3, whilst red free energy peaks correspond to the QLL/ice interface defined as the outermost peak $> \frac{1}{2}$ k$_B$T

density peaks coloured dark blue are the outermost ice bilayers and the interfaces lie at the density minima between the two regions.

Our second approach defines the QLL/ice interface as the outermost ΔF_z^* free energy barrier greater than $\frac{1}{2}\ k_BT$. The energy peaks identified with this criterion are coloured red in Figure 7.1, and the QLL/ice interfaces lie at the corresponding maxima. This criterion recognises that the QLL/ice interface is characterised by a large diffusion free energy barrier whereas diffusion barriers within the QLL are comparatively weak. The equipartition theorem states that the energy associated with each quadratic degree of freedom of a molecule is $\frac{1}{2}\ k_BT$. A diffusing molecule in thermal equilibrium has an average translational kinetic energy of $\frac{1}{2}\ k_BT$ along the z-axis, and therefore a barrier $\Delta F_z^* > \frac{1}{2}\ k_BT$ hinders diffusion.

Heat maps of $\cos\beta$ and γ layer-by-layer distributions are identified in Chapter 4 as

a useful tool for qualitatively illustrating the disorder/order transition at the QLL/ice interface. Here we use them to assess the positioning of the QLL/ice interface using our two criteria. Figure 7.2 shows the layer-by-layer orientation heat maps for the basal and prism TIP4P/Ice surfaces at the seven simulated temperatures. Traversing the figure from left to right corresponds to density slices deeper and deeper into the bulk. The positions of the interfaces from the $\rho(z)$ and ΔF_z^* criteria are indicated by cyan and red lines, respectively. The red and cyan lines demonstrate that the two criteria successfully separate the disordered and ordered bilayers and provide a reliable description of the QLL/ice interface position. At low temperatures the

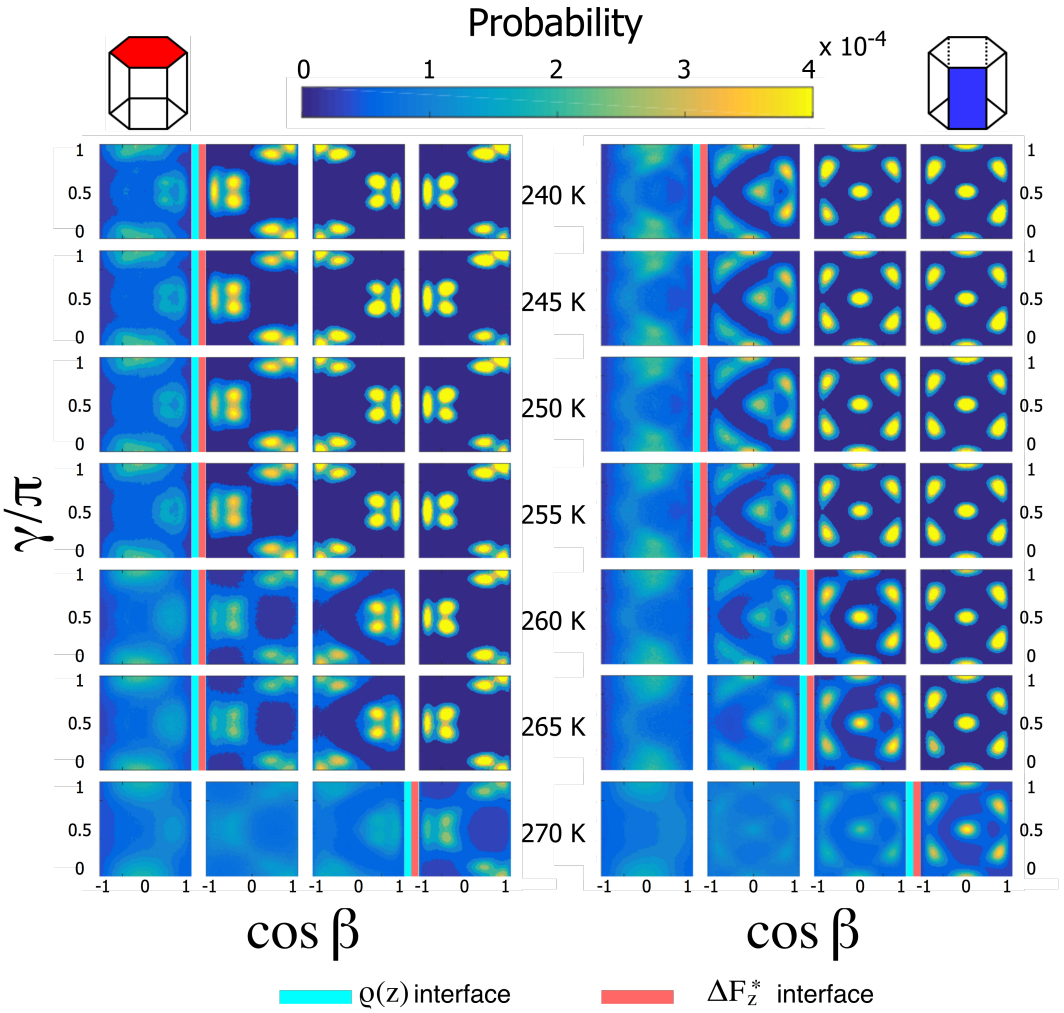


Figure 7.2: Orientation heat maps of water molecules within slices corresponding to $\rho(z)$ peaks for the (left) basal and (right) prism TIP4P/Ice surfaces. Traversing left to right corresponds to density slices deeper and deeper into the bulk. The interface defined using our $\rho(z)$ criterion is marked by an cyan line. The interface defined using the ΔF_z^* criterion is indicated by red lines.

criteria suggest that the QLL corresponds to the first density peak and is therefore composed of one bilayer of molecules. This is consistent with the consensus from MD simulations using a range of different water models and further validates our criteria [7, 36–38].

At 270 K, the equilibrium TIP4P/Ice QLL size fluctuates with large amplitudes. Analysis of the liquid-like cluster sizes in Chapter 4 suggests that the equilibrium QLL at 270 K is dynamically melting and crystallising ice and QLL bilayers. The QLL/ice interface position at 270 K therefore fluctuates and does not remain between two fixed bilayers. To ensure an accurate description of the QLL/ice interface position, at 270 K we split our simulations into 10 equally sized sections of around 40 ns and determine the interface position independently for each decile. The basal and prism interface positions shown in Figure 7.2 at 270 K are averaged over the splits.

ΔF_z^* exhibits distinct diffusion free energy barriers at the basal and prism QLL/ice interfaces. The influence of the barriers on the kinetics at the QLL/ice interface is captured by monitoring the rate at which molecules diffuse across the interface during our TIP4P/Ice MD simulations. A molecule is identified as having crossed the interface if the z-coordinate of its oxygen atom passes through the z-coordinate of the interface. Crossings are only counted for molecules which do not return back for at least 1.0 ns to ensure that we do not count molecules which may be vibrating near the interface.

The rates of molecular diffusion from the QLL into ice and from ice into the QLL are plotted in Figure 7.3. The top row is the result from the $\rho(z)$ interface criterion and the bottom row is the result for the ΔF_z^* interface criterion. The simulations are split into ten equal slices and the rate of crossings is determined independently for each decile and for both surfaces within each slab. The mean value obtained from the two slab surfaces and ten splits is plotted in Figure 7.3 and the error bars indicate the standard deviation of the mean. The results show that the rate of molecules crossing into the QLL and out of the QLL are equivalent. The QLL is at equilibrium and for each molecule leaving the QLL to join ice, a molecule joins the QLL from ice. Previous MD studies of the basal and prism surfaces of ice at 250 K using a

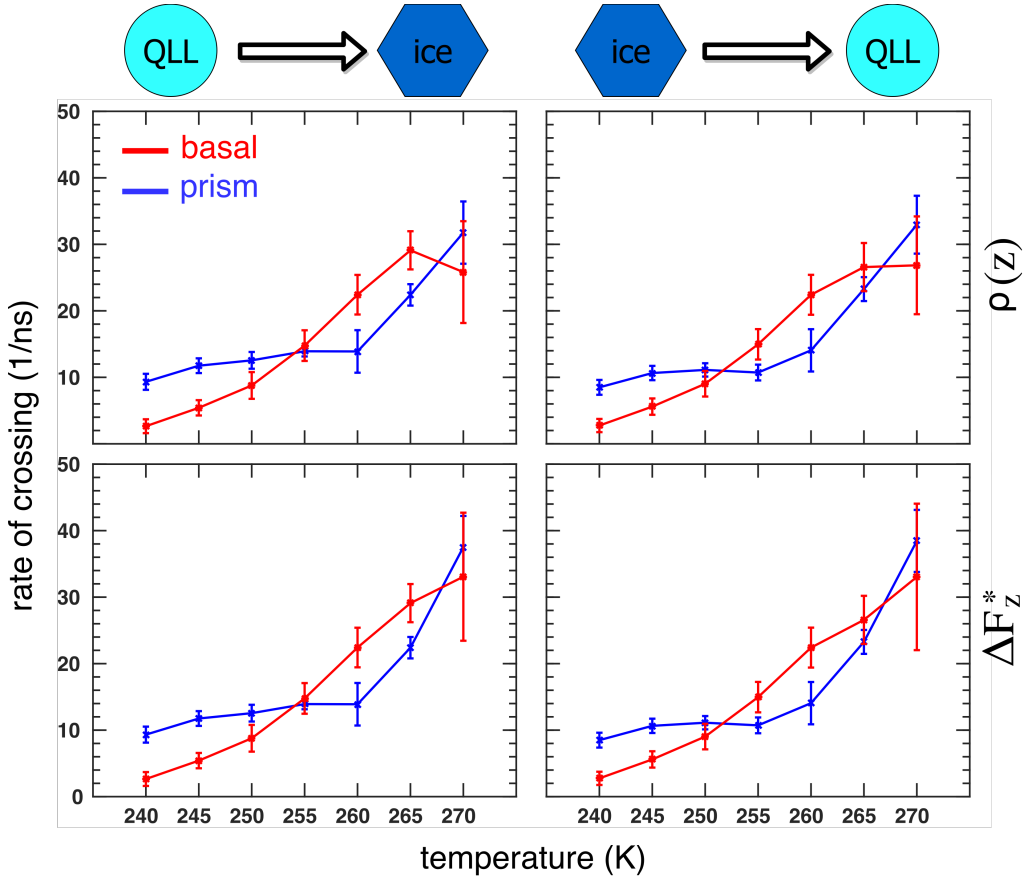


Figure 7.3: Rate of crossings at the QLL/ice interface into the ice sublayer (left) and into the QLL (right). Top row: interface position determined in relation to the innermost density peak with over 75% liquid-like molecules. Bottom row: interface position defined as the outermost diffusion barrier greater than $\frac{1}{2}\ k_B T$

six-site water model report that transitions between the QLL and ice occur as binary exchanges, in agreement with our findings [29, 38]. The two interface criteria produce a consistent result. As temperature increases, the rates of crossing increase and the basal and prism rates cross-over twice.

7.1.1 Discussion

The rates of QLL/ice interface crossings provide a measure of the kinetics at the interface. The basal and prism rates are distinct and reveal two cross-overs which correlate with the two cross-overs in the morphology diagram described in detail in Section 1.1. At low temperatures, the rate of diffusion across the prism interface is fastest, implying that thin plates form. At intermediary temperatures, the rate of

diffusion across the basal interface is fastest, implying that elongated prisms form. Close to the melting point, the rate of diffusion across the prism interface is fastest again, implying that thin plates form. This result leads us to the hypothesis that the growth shape of I_h crystals is related to the rate of QLL/ice molecular crossings and the underlying diffusion energy barriers.

Ice growth in vapour involves the crystallisation of a QLL in equilibrium with ice and vapour. The vapour phase is the source of QLL molecules and fuels QLL growth. The ice phase plays a key role and is the limiting factor for determining the relative rate of growth of the two surfaces. The implication of our results is that the QLL/vapour interface plays a secondary role in the overall kinetics of vapour growth and that the rate of vapour adsorption onto the basal and prism surfaces are very similar. To test our hypothesis we perform metadynamics simulations to recover the energetics of adsorption and desorption at the QLL/vapour interface.

7.2 QLL/Vapour Interface

Metadynamics simulations are performed to uncover the energies associated with the binding and desorption of vapour molecules onto the surfaces of ice. The simulation protocol is described in detail in Section 3.5.2. Simulations of the basal and prism surfaces at 260 K are performed using TIP4P/Ice. Only one temperature is simulated due to the high computational cost of metadynamics. During the metadynamics simulations, the z-coordinate of a selected oxygen atom is biased. The result is the enhanced sampling of diffusion of the selected molecule through vacuum, the QLL and ice, enabling the energy cost associated with diffusion, adsorption and desorption to be recovered. Snapshots of the simulations are shown in Figure 7.4. The biased atom is coloured in blue. During the simulations the biased molecule explores the region outlined by coloured atoms. Energy boundaries restrict the CV and prevent the molecule from exploring the regions corresponding to the silver atoms. This restricts the CV space that the molecule can explore resulting in faster convergence of the free energy within the allowed CV space.

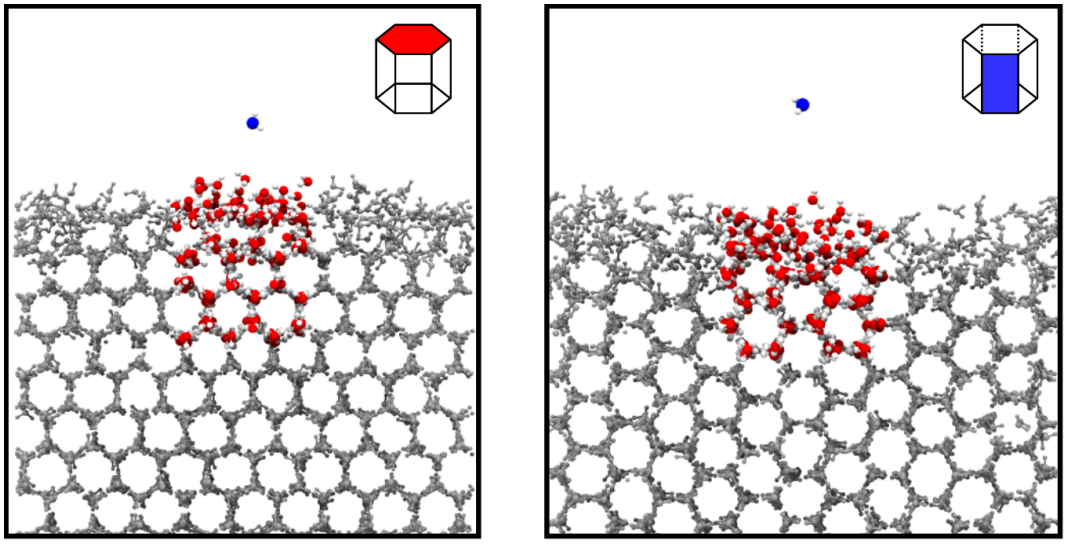


Figure 7.4: Snapshot of the basal and prism metadynamics simulations at 260 K using TIP4P/Ice. The selected biased atom is shown in blue exploring the vacuum.

Metadynamics simulations of the TIP4P/Ice basal and prism surfaces are performed for 1100 ns at 260 K. Figure 7.5 shows the z-coordinate CV during the simulations. When the z-coordinate is above around 90 Å, the molecule is in the vacuum region where its mobility is drastically increased. In this region, the z-coordinate fluctuates by large amounts. At low z-coordinates in the scale, the molecule is in the bulk and is subject to restricted mobility due to the rigid and stable bonding network. Within the bulk, the energy barriers of diffusion are much greater and characterised by deep free energy basins. The biased molecule therefore spends more time in a given bulk state as more repulsive Gaussians are required to fill the free energy well and to enable the molecule to escape and explore neighbouring bilayers.

The free energy profile is reconstructed from the metadynamics simulations and plotted along side the CV in Figure 7.5. The basal and prism profiles are plotted over the range of coordinates that are well converged. The water molecule is in the vapour phase above around 90 Å along the profiles. Our results show that free energy of a vapour molecule converges to an identical value for the basal and prism surface as observed for z-coordinates above around 96 Å. This result is expected because the vapour phase is modelled as vacuum and there are no neighbouring molecules that influence the potential energy of the vapour molecule. The implication of our

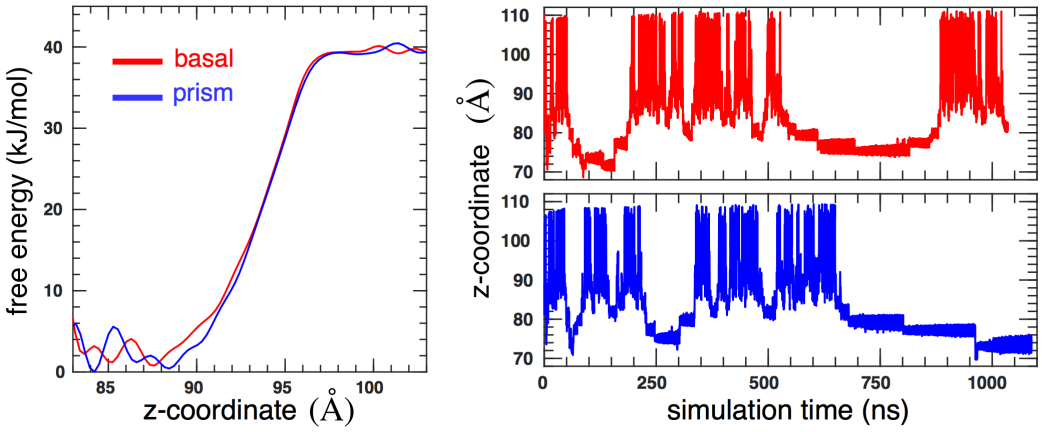


Figure 7.5: Free energy associated with a molecule adsorbing onto and diffusing through the basal and prism surfaces. The profiles are recovered from metadynamics simulations of the two surfaces using TIP4P/Ice. The CV is the z-coordinate of a selected molecule and this is shown in the figures on the right for the two surfaces.

result is that the interactions between TIP4P/Ice molecules have a maximum range of approximately 6 Å since the QLL/vacuum interface occurs at $z \approx 90$ Å. The first QLL sublayer, at the interface with vacuum, occurs at $z \approx 87.6$ Å and 88.5 Å for the basal and prism surfaces, respectively. The neighbouring minimum along the basal profile at around 85.3 Å is the first ice sublayer at the QLL/ice interface. The minimum at 86.6 Å along the prism profile corresponds to a second QLL sublayer based on our $\rho(z)$ and ΔF_z^* criteria.

The convergence of our simulations is monitored by calculating ΔF_{ab}, the difference in free energy between two states a and b. The free energy of the molecule in the vapour state, F_{vap} is calculated as the mean free energy between 99 Å and 104 Å, and between 98 Å and 102 Å, for the basal and prism surfaces, respectively. The free energy of the molecule in the first QLL sublayer, at the interface with vacuum, F_1 corresponds to the free energy minimum between 86.5 Å and 88.5 Å, and between 87.8 Å and 89.3 Å, for the basal and prism surfaces, respectively. The free energy of the molecule in second bilayer, F_2 corresponds to the free energy minimum between 84.5 Å and 86 Å, and between 85.8 Å and 87.3 Å, for the basal and prism surfaces, respectively. The convergence is computed from the three states as $\Delta F_{v1} = F_{vap} - F1$, $\Delta F_{v2} = F_{vac} - F2$ and $\Delta F_{12} = F_1 - F2$ by reconstructing the free energy profile repeatedly every time 100 Gaussians are deposited. The results are plotted in Figure

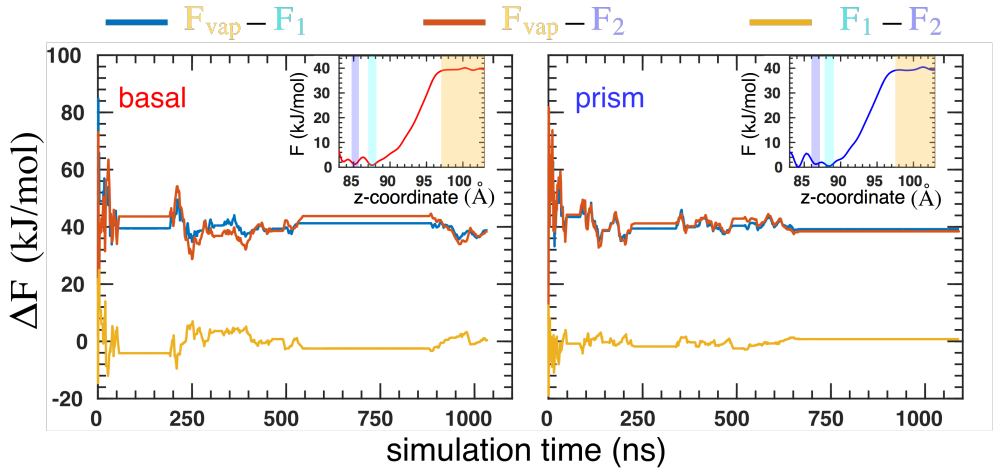


Figure 7.6: Convergence ΔF of the basal and prism metadynamics simulations biasing the z-coordinate of a molecule. The three regions used to compute ΔF are highlighted in the free energy profiles in the inset. The mean free energy of the vapour phase, F_{vap} is calculated from the yellow region, whilst the free energy minima within the first and second bilayers, F_1 and F_2, are calculated from the minima in the cyan and dark blue regions, respectively.

7.6. The insets in the plots indicate the position of the first and second free energy minima used to calculate the convergence in light and dark blue, respectively, as well as the vacuum region in yellow. The convergence in Figure 7.6 is constant during the time steps where the molecule is in deeper bulk sublayers. During these periods, no Gaussians are deposited within the region that the convergence is monitored and therefore the difference in free energies remain constant. Nonetheless, the results are reasonably well converged and the amplitudes of the fluctuations decrease with time. ΔF_{12}, the difference in free energy between the first and second bilayers is converged to approximately zero for the two surfaces, whilst ΔF_{v1} is converged to a value just under 40 kJ/mol for the two surfaces.

The free energy barriers associated with diffusion at the QLL/ice interface, ΔF^*_{meta} are recovered from our metadynamics simulations. At 260 K, ΔF^*_{meta} is 3.3 kJ/mol and 4.4 kJ/mol for the basal and prism surfaces, respectively. This barrier is previously estimated from ΔF^*_z defined in relation to the logarithm of $\rho(z)$. At 260 K, ΔF^*_z yields 3.9 kJ/mol and 4.9 kJ/mol for barriers at the basal and prism QLL/ice interfaces, respectively. ΔF^*_z from MD and ΔF^*_{meta} from metadynamics corroborate that the prism QLL/ice diffusion barrier is greater than the basal barrier. Experimental studies using an environmental molecular beam method measure the

activation energy associated with the incorporation of a QLL molecule into ice at 6 ± 5 kJ/mol [131]. This is in good agreement with our metadynamics results and consistent with our MD ΔF_z^* barriers within the margin of error. An x-ray adsorption spectroscopy study of supercooled and normal liquid water between 251 K and 288 K suggests that the activation energy required to break a hydrogen bond, whilst leaving all the molecules in essentially the same position, is 6.3 kJ/mol [132]. Our results therefore indicate that the diffusion energy barrier at the QLL/ice interface corresponds to the activation energy required to break a hydrogen bond. This is consistent with observations that a water molecule forms four hydrogen bonds in ice bilayers and three hydrogen bonds in the QLL sublayer at the QLL/ice interface [29, 38].

The Arrhenius equation gives the dependence of the rate constant, k on temperature and activation energy, E_a. It is defined as

$$k = Ae^{-E_a/(RT)} \tag{7.2}$$

where A is the pre-exponential factor and R is the gas constant. The ratio of the rate constants for diffusion across the basal and prism QLL/ice interfaces can be expressed as

$$\frac{k_{basal}}{k_{prism}} = \frac{e^{-E_b/(RT)}}{e^{-E_p/(RT)}} \tag{7.3}$$

where E_b and E_p are the activation free energies for diffusion across the basal and prism interfaces, respectively. The activation free energies can be equated to the diffusion energy barriers estimated from ΔF_z^* and ΔF_{meta}^*. The value of the ratio is large when the basal rate of diffusion is fastest and small when the prism rate of diffusion is fastest. The ratio is calculated from our estimates of ΔF_z^* at the seven simulated temperatures and from our ΔF_{meta}^* estimated from metadynamics at 260 K, and plotted in Figure 7.7.

The rate constant ratios from ΔF_z^* reveal that below $\sim$ 257.5 K and above $\sim$ 267.5 K the rate of diffusion at the prism QLL/ice interface is fastest whilst at the intermediary temperatures the basal rate of diffusion is fastest. This result correlates

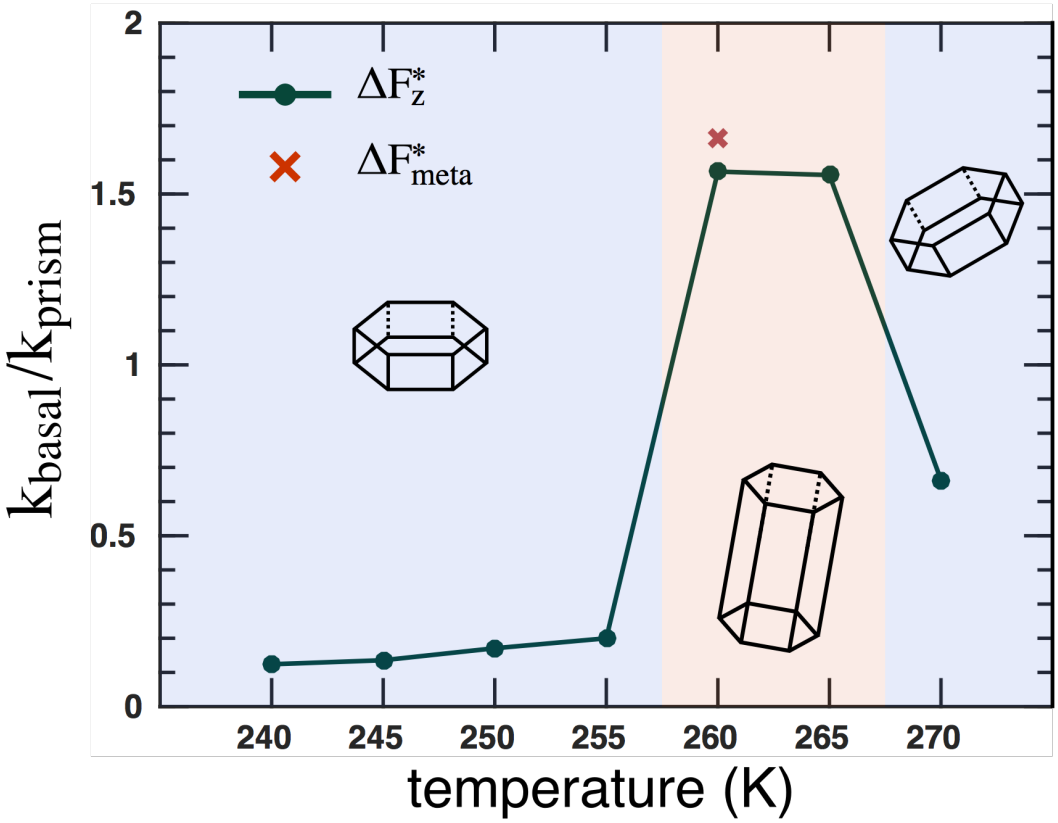


Figure 7.7: Ratio of the basal and prism rate constants associated with diffusion at the QLL/ice interface, estimated from ΔF_z^* and ΔF_{meta}^*. Thin plates, which occur when the prism surface grows fastest, are predicted when the prism rate constant is greater than the basal rate constant. Elongated prisms are predicted when the basal rate constant is greater.

with the rate of crossings results, which are an indirect measure of the same diffusion energy barriers. The agreement of the ratio of the rate constants with the morphology diagram of ice suggests that the two cross-overs from thin plates to elongated prisms are due to the diffusion energy barriers at the QLL/ice interface. At 260 K, the ratio from ΔF_z^* and ΔF_{meta}^* yield a reasonably consistent result and validates the robustness of the ΔF_z^* approach based on the logarithm of $\rho(z)$.

Whilst $\Delta F(z)$ enables the diffusion free energy barriers, ΔF_z^* within ice and the QLL to be estimated from MD simulations, the free energy cost of desorption cannot be recovered using this approach. The metadynamics simulations bias the z-coordinate of a selected molecule resulting in the sampling of many adsorption and desorption events. The activation energy of desorption is estimated from metadynamics as 38.8 kJ/mol and 39.2 kJ/mol at the basal and prism surfaces,

respectively. This indicates that the basal desorption activation energy is only 1% smaller than the prism activation barrier, and therefore desorption is only very weakly dependent on the crystallographic surface. Our previous MD simulations show that the QLL/ice interface is key to the anisotropy of the basal and prism growth rates and imply that the QLL/vapour interface does not impact the relative rates. Analysis of the dangling bond structure in Chapter 4 suggests that the basal and prism QLL surface structures are closely related. These metadynamics results at 260 K confirm these findings.

The hydrogen bond energy is approximately 23 kJ/mol [117]. Previous MD studies show that surface QLL molecules at the QLL/vapour interface form two hydrogen bonds [29, 38]. Breaking the two hydrogen bonds therefore requires 46 kJ/mol. This is comparable to our metadynamics estimates of the desorption activation energies of around 39 kJ/mol. Our results therefore indicate that desorption involves breaking two hydrogen bonds and enable us to estimate a hydrogen bond strength of 20 kJ/mol for the TIP4P/ice water model at 260 K. Our estimate of the desorption activation energy is lower than the experimental enthalpy of sublimation of 51.6 kJ/mol at 260 K [133]. However, previous experimental studies of ice surfaces between 170 K and 200 K measure the activation energy of desorption at 42 kJ/mol, in close agreement with our findings [131]. Whilst desorption is an activated process, there are no energy barriers hindering adsorption. Our metadynamics results show that adsorption is a favourable process with no activation energy. A vapour molecule adsorbing onto the basal and prism QLL surfaces is therefore not subject to surface specific barriers.

7.2.1 Discussion

Our metadynamics results show that the desorption free energy costs are equivalent for the basal and prism surfaces. This is consistent with our MD rate of crossing results which imply that the QLL/vapour interface does not introduce anisotropy into the basal and prism growth kinetics. Adsorption is a barrierless and favourable process. The surface accommodation coefficient, α_s provides a measure of the stickiness of a surface and can be used to explain our metadynamics results.

The surface accommodation coefficient, α_s is defined as the ratio of the number of water vapour molecules that adsorb onto an ice surface to the total number of vapour molecules striking the surface [29, 38]. A value of $\alpha_s = 1$ indicates that every incident vapour molecule adsorbs onto the surface. MD simulations of water surfaces indicate that $\alpha_s \approx 1$ at 300 K [134]. Experimental and simulation studies of ice surfaces between 20 K and 120 K also report $\alpha_s \approx 1$ [135]. Only when the grazing angles and kinetic energy of incident vapour molecules is large, is a value of 0.9 measured. MD simulations of the basal surface at 250 K estimate $\alpha_s > 0.99$ [29]. Experimental and simulation works affirm that α_s is close to 1 for both ice and water surfaces. This implies that the interaction of incident vapour molecules with ice surfaces are overwhelmingly favourable and independent of the crystallographic plane. This is consistent with our metadynamics results which reveal that there are no activation barriers associated with adsorption of a vapour molecule onto the basal and prism surfaces of ice.

The incorporation lifetime, t_i of QLL molecules before they are incorporated into ice is calculated from previous MD simulations for the basal and prism surfaces using the 6-site NE6 water model [29, 38]. A lifetime of 700 ns and 200 ns is estimated for the basal and prism surfaces, respectively, at a simulated temperature of 250 K which corresponds to $-40°$C for the NE6 water model used. A long incorporation lifetime implies a slow rate of QLL/ice crossing. At low temperatures, around $-40°$C, our MD rate of crossing results show that the basal rate of QLL/ice crossings is slower than the prism rate, consistent with the lifetime estimates. The vapourisation lifetime, t_v of QLL molecules before they join the vapour phase is reportedly 1200 ns and 1400 ns for the basal and prism surfaces, respectively, based on MD simulations at $-40°$C [29, 38]. In contrast to the t_i, t_v is not strongly dependent on the crystallographic surface, consistent with our metadynamics desorption activation energies.

Whilst the surface accommodation coefficient of ice surfaces is close to 1 and does not display any strong dependence on the crystallographic surface, the bulk accommodation coefficient, α_b of vapour molecules into ice is complex and uncertain and experimental measurements vary by more than 2 orders of magnitude [131, 136–

138]. A recent study suggests that the bulk accommodation coefficient decreases from close to unity at 170 K to around 0.41 at 200 K [131]. The accommodation coefficient into the bulk depends on a competition between of the rate of QLL to vapour desorption and QLL to ice incorporation. The rate constant for incorporation into the bulk is found to increase with temperature, consistent with our rate of crossing results, however the experimental measurements are not sensitive to the crystallographic surface and incapable of detecting any cross-overs for the basal and prism surfaces [131]. Experimental results suggest that the rates of desorption and incorporation are comparable at 200 K but at lower temperatures the rate of incorporation is greater [131]. α_b can be defined in terms of the vapourisation and incorporation lifetimes and α_s as

$$\alpha_b = \alpha_s \frac{t_i^{-1}}{t_v^{-1} + t_i^{-1}} \tag{7.4}$$

Assuming $\alpha_s \approx 1$, the bulk accommodation coefficient for incorporation of a vapour molecule into ice is estimated from MD simulations at 0.63 and 0.87 for the basal and prism surfaces at $-40°$C, respectively, using the 6-site NE6 water model [29, 38]. α_b takes into account both the QLL/ice and QLL/vapour interface and yet the result is consistent with our rate of crossing findings which only takes into account the QLL/ice interface. The implication is that the QLL/ice interface is rate limiting which is in agreement with recent commentary by Neshyba [130] based on MD estimates of the bulk accommodation coefficient.

One limitation of our rate of crossings approach is the use of a rigid water model which does not allow OH covalent bonds to be broken or formed. In reality the hydrogen atoms that an oxygen atom is covalently bonded to constantly change with the motion and dynamics of the atoms. Hydrogen bonded atoms may become covalently bonded and vice versa. We propose, based on our MD and metadynamics simulation results, that the source of the anisotropy of ice crystals is the kinetics at the QLL/ice interface. The dynamics of an oxygen atom crossing the interface is determined, in part, by hydrogen bonds. It is reasonable to assume that a molecule crossing the interface becomes hydrogen bonded to molecules in the neighbouring

interface during the crossing process. The mechanism could involve a pulling effect by the attractive hydrogen bonds. During crossing, the oxygen atom may become unbonded to its own hydrogen atoms and transition from hydrogen bonded to covalently bonded to hydrogens in the approaching bilayer. This remains speculative until simulations with non-rigid bonds are performed. Further experimental and simulation studies of the anisotropy of hydrogen bonds at the QLL/ice interface will provide valuable insight into the mechanism of ice growth, whilst nuclear quantum effects remain to be investigated.

The QLL/ice interface plays a key role in the kinetics of ice growth. Whilst the exact mechanism of crystallisation remains unknown, we may speculate on some of the possible underlying processes occurring at the interface and how to detect them. The interface may provide a cooling effect on the QLL by removing high kinetic energy QLL molecules and replacing them with low kinetic energy ice molecules. This could be probed by studying the kinetic energy of the crossing molecules. Alternatively, the exchange of molecules between the QLL and ice may catalyse the QLL ordering by providing the QLL with ice molecules which are already in low energy ordered configurations. A previous MD study of the adsorption of vapour molecules onto the basal surface notes that the vapour molecules which successfully absorb into the inner QLL sublayer are initially in orientations that complement existing vacancies within the sublayer [29]. A similar analysis at the QLL/ice interface could help uncover the mechanisms of QLL crystallisation.

Chapter 8

Deep neural network for identifying ice and water molecules

Identifying and classifying the structural phase of molecules is a key task in computer simulations. Studies involving phase transitions rely on the accurate classification of crystal structures whilst simulation studies of nucleation depend on an accurate distinction of the liquid from the crystal. Investigating the formation of defects in crystals involves their detection using an automated and reliable process. Simulation studies of ice surfaces at temperatures above the onset of surface melting rely on being able to accurately distinguish between ice and water-like QLL molecules using an automated process. Throughout this thesis we have used the q_3 parameter, described in Section 3.4.1 to identify ice and QLL molecules. In this section we present a novel approach for identifying molecules using a deep neural network which achieves a remarkable accuracy with an error rate over ten times smaller than the q_3 method. Our work demonstrates that a trained neural network can accurately identify bulk and surface water and ice molecules using a flexible framework which does not rely on any intuition on the underlying structure, and can be generalised to any phase. Our approach builds upon the neural network developed by Geiger [139] which uses as input a collection of symmetry functions that are sensitive to the

positions of atoms. We propose a neural network which simply requires as input the atomic coordinates of the molecules in the system, and can adapt to any symmetry without modification.

8.1 Steinhardt parameter, q3

An overview of the Steinhardt method is discussed in Section 3.4.1. In previous chapters q_3 is used as shorthand for $\bar{q}_3$, defined in Eq. 3.4. In this chapter q_3 refers to the original Steinhardt definition defined in Eq. 3.3, and $\bar{q}_3$ refers to the modification by Lechner [101] defined in Eq. 3.4. We introduce a further modification of q_3, $\tilde{q}_3$ which we develop in an effort to improve the accuracy of $\bar{q}_3$. $\tilde{q}_3$ is defined for molecule i as

$$\tilde{q}_3(i) = \frac{1}{N_b(i)+1} \sum_{j=0}^{N_b(i)} q_3(j) \tag{8.1}$$

where $N_b(i)$ is the number of nearest neighbours within the cut-off distance of 3.5 Å and q_3 is averaged with its nearest neighbours and itself.

One of the disadvantages of the Steinhardt approach is that the choice of spherical harmonics relies on intuition and advanced knowledge of the underlying structures. Steinhardt bond order parameters are poor at distinguishing phases which contain multiple distinct local spatial environments within each phase [139]. In practice choosing the ideal spherical harmonics for identifying many different phases involves a lengthy trial and error process. The $l = 3$ spherical harmonics are an appropriate choice for distinguishing water and hexagonal ice molecules, however, it is a poor choice for distinguishing water and cubic ice molecules. Motivated by a desire to simplify the process and improve the accuracy, we design a neural network to identify ice and water molecules which can be generalised to phases where the underlying structure is unknown and where multiple bonding environments are present. It has previously been shown that a neural network trained using a set of symmetry functions can be used to distinguish between liquid water and several phases of ice over a range of temperatures and pressures [139]. Our neural network demonstrates that

the phase of a molecule can be classified using as input simply the xyz-coordinates of its nearest neighbours without compromising the accuracy.

8.2 Artificial neural network

An artificial neural network is a computational approach inspired by the brain. The network consists of layers of interconnected computational units known as neurons. Typically, a neuron receives and processes an input by passing it through a non-linear activation or transfer function. Once processed, the neuron feeds the output into the next layer in the network. The neurons in neighbouring layers are connected via synapses which store weights that manipulate the inputs and outputs. The neural network used in this work is represented in Figure 8.1. Black lines represent the weights. These weights, along with the biases associated with each neuron, are adjusted during the training of the network until the desired outcome is reached.

Neural networks are composed of three parts. The first is the input neuron layer which receives external information. The input layer feeds into the second part which is then composed of one or more computation layers known as hidden layers. The hidden layers process the input signals with weights and biases, and pass their output into the final output layer. The output layer computes the final predictions and converts the output into the desired form.

The topology of our neural network with two hidden layers is described by $I \times H \times H \times O$, where I, H and O are the number of nodes in the input, hidden and output layers respectively. An input X is passed into the neurons in the first hidden layer as a weighted sum over all the input neurons. The signal into neuron m in the first hidden layer is

$$\sum_l w_{ml}^{(1)} X_l \tag{8.2}$$

where $w_{ml}^{(1)}$ is the synaptic weight connecting the input neuron l with the neuron m in the first hidden layer. The signal into each neuron is then shifted by a bias

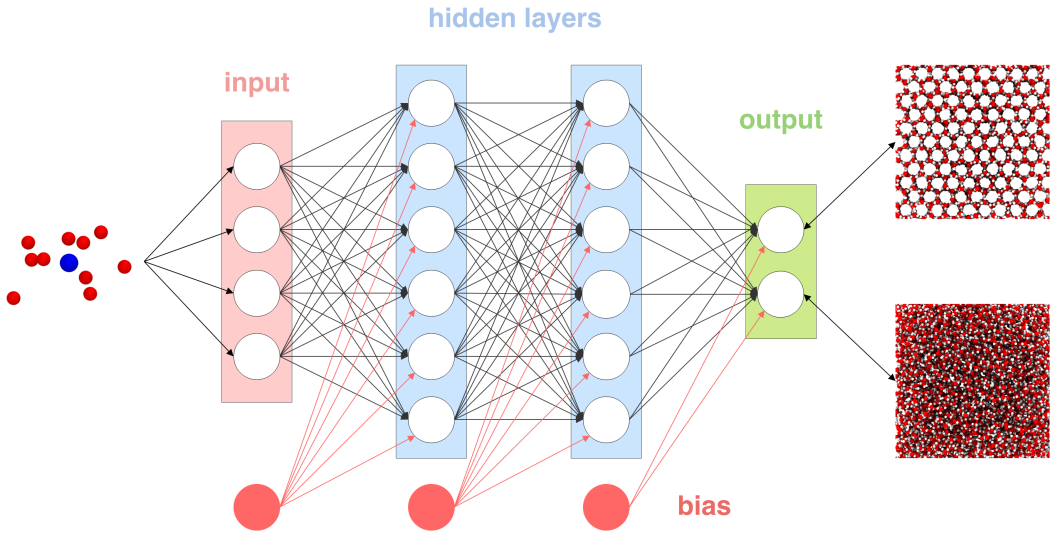


Figure 8.1: Schematic representation of a neural network with two hidden layers. Black lines correspond to weights and red lines correspond to biases which are tuned during training. The input is the coordinates of the nearest neighbours to a molecule and the output is the predicted phase.

$b_m^{(1)}$ and processed by an activation function, $\phi(x)$. The overall output, known as activation, of the neurons in the first hidden layer is

$$a_m^{(1)} = \phi\left(\sum_{l=1}^{I} w_{ml}^{(1)} X_l + b_m^{(1)}\right) \tag{8.3}$$

The activations of the first hidden layer, $a^{(1)}$ are then inputted into the second hidden layer using the same process. The activation of the neuron n in the second hidden layer is

$$a_n^{(2)} = \phi\left(\sum_{m=1}^{H} w_{nm}^{(2)} a_m^{(1)} + b_n^{(2)}\right) \tag{8.4}$$

In the last step, the activations of the second hidden layer are passed into the output layer where the signals are weighted, biased and processed by an activation function $\tilde{\phi}$. The final product of the output neuron p is

$$y_p = \tilde{\phi}\left(\sum_{n=1}^{H} w_{pn}^{(3)} a_n^{(2)} + b_p^{(3)}\right) \tag{8.5}$$

The activation function of the hidden units is the rectified linear unit, $\phi(x) = \max(0, x)$, whilst a linear activation function is used for the output layer, $\tilde{\phi}(x) = x$.

The output y of the network is a vector with length equal to the number of structures the network is trained to detect. The sum of the y elements is equal to one and individually each element gives the probability that the input is a particular structure. An input is then classified as the most probable class. The activation functions and large number of weights linking the interconnected neurons results in a complex non-linear relationship between the input and output. The network is trained by providing it with a set of atomic configurations with known structures, and monitoring the classification error. The framework is flexible and during the training process the weights and biases are adjusted resulting in an accurate structure recognition network.

8.2.1 Nearest Neighbours Input

The input of the neural network developed by Geiger [139] is, for a given atom, a set of symmetry functions which depend on the Cartesian coordinates of its neighbours. The choice of symmetry functions is key to the accuracy of their neural network and must be sensitive to the local environments of the atoms. The process of selecting the symmetry functions involves first outlining a set of functions believed to be appropriate for the given structures. This requires initially studying the distribution of distances and angles within the local environments to gain insight into the structural features that can be used to distinguish between the different phases. Based on these considerations, candidate symmetry functions are defined and their distributions within the different phases are computed. The symmetry functions with the smallest overlap are then shortlisted. The next step involves a sensitivity analysis whereby the neural network is trained with the selected symmetry functions and any function which contributes weakly to the predictions is removed. The final outcome is 30-40 symmetry functions which are then used to retrain the neural network. If a new structure is introduced, the entire process must be repeated.

Our neural network does not require the lengthy process of shortlisting and selecting symmetry functions and sensitivity analysis. Our input for a given atom is simply the xyz-coordinates of its nearest neighbours relative to itself. These are

computed for atom i by subtracting its coordinates from the xyz-coordinates of its neighbours. The simplicity does not compromise the accuracy and the only parameter of our input is the number of nearest neighbours. Water has 4 nearest neighbours in the first coordination shell. We use the 10 nearest-neighbours for our input, to take into account the second coordination shell which is composed of approximately 14 molecules in water and ice. Our input layer is therefore composed of 30 input neurons.

8.2.2 Training the neural network

The training process involves tuning the weights and biases which are depicted as lines in Figure 8.1. The training set $\mathcal{T}$ consists of the nearest neighbours coordinate vector $\mathbf{X}$ and the corresponding structure vector $\tilde{\mathbf{y}}$,

$$\mathcal{T} = \{\mathbf{X}, \tilde{\mathbf{y}}\} \tag{8.6}$$

$\tilde{\mathbf{y}}$ is the output vector from the neural network and encodes the classes of the structures. In the case of classifying ice and water, $\tilde{\mathbf{y}}^{(i)}$ is comprised of two elements for each entry i. If the structure of entry i is hexagonal ice, $\tilde{\mathbf{y}}^{(i)} = \{1, 0\}$, and if it is water $\tilde{\mathbf{y}}^{(i)} = \{0, 1\}$.

The training set is a set of relative coordinates of bulk and liquid water molecules. The training set is produced by running TIP4P/Ice MD simulations of the phases in question. For this study, simulations of I_h and supercooled water at 260 K are performed. Each system is composed of 2880 molecules and periodic boundary conditions ensure that there are no interfaces or surfaces. The TIP4P/Ice simulation protocol described in Section 3.3 is used and the systems are equilibrated for 0.24 ns. Production runs are performed for 1 ns in the NVT ensemble and 1001 trajectory frames are saved every 1000 fs. The nearest neighbours are collected for each molecule producing a training set with over 5×10^6 entries and a validation set nearly 6×10^5 configurations in total. Trajectory frames are collected every 1000 fs to avoid correlations between entries in the training set, whilst 1 ns of simulations is sufficient to ensure that thermal fluctuations in the structure are captured. When

calculating the relative coordinates of the nearest neighbours, the y-dimensions of the simulation boxes are increased by 30 Å to create the effect of an ice/vacuum and water/vacuum interface. This enables the neural network to learn to identify surface ice and water molecules as well as bulk ice and water molecules. It is assumed that there are no phase transitions during the simulations which is verified by visual inspection. In general, monitoring the energies during the simulation can indicate whether any structural changes have occurred. If a phase transition does occur, a revised simulation protocol where no transitions occur must be found and the parts of the simulation with the new phase must not be included for training.

The classification error of the network is monitored during the training procedure and used to drive the tuning of the weights and biases. The cross-entropy loss function is used by our neural network and is defined as

$$E = -\frac{1}{n_t}\sum_{i}^{n_t}\left[\tilde{y}_t^{(i)}\log\left(y_p(X^{(i)})\right) + \left(1-\tilde{y}_t^{(i)}\right)\log\left(1-y_p(X^{(i)})\right)\right] \tag{8.7}$$

where n_t is the total number of items in the training data, $\tilde{y}_t^{(i)}$ is the desired output for entry i and $y_p(X^{(i)})$ is the predicted output. E is minimised using the back-propagation algorithm [140].

The network is trained using a first-order gradient-based optimiser known as Adam [141] with a learning rate of 0.001. Mini-batch learning is performed using a batch size of 2000. The weights and biases are tuned each time the learning procedure cycles through every entry in the batch. Each cycle through the entire data set is known as an epoch and training is allowed for a maximum of 5000 epochs. The weights and biases are initially randomised and at the beginning of each epoch the order of training data is randomised to avoid correlations. During training the proportion of incorrect classifications within the validation set is monitored to ensure that the network does not overfit. Overfitting occurs if the validation error increases whilst the training error continues to decrease. It is avoided using early stopping whereby the training procedure is halted if the validation error increases during six

consecutive epochs.

The overall architecture of our neural network is $30 \times 256 \times 256 \times 2$. There are 256 neurons in each hidden layer and the output detects two phases, water (QLL) or I_h. The training process takes a few days using 4 CPUs on a conventional desktop and 74242 parameters are optimised. The ten nearest neighbours are used to produce the input, which takes into account the first neighbour shell and nearly half of the second neighbour shell. It is anticipated that the larger the number of nearest neighbours included in the input, the higher the accuracy of the network. The neural network developed by Geiger [139] showed an error of 85% for ice V and 98% for ice III when including two neighbour shells and an increase to nearly 100% when including three neighbour shells.

8.2.3 Classification Errors

In order to estimate the error of the network we run six further independent MD simulations of ice and water at 240 K, 245 K, 250 K, 255 K, 265 K and 270 K using the protocol described previously. The test set is thus obtained from simulations that are independent from the training set and at different temperatures. The nearest neighbours are calculated without extending the y-dimension so that no surface or interface is included. The six water and ice simulations are concatenated into one

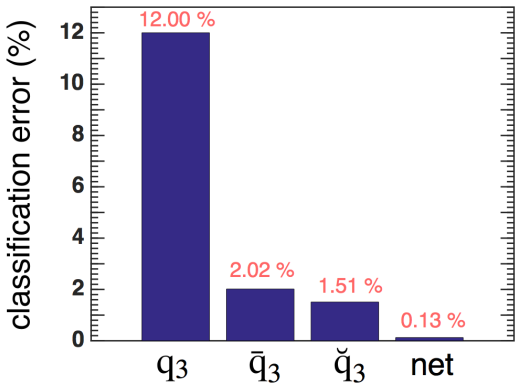


Figure 8.2: Classification error of the Steinhardt and neural network algorithms determined from six simulations of ice and water at 240 K, 245 K, 250 K, 265 K and 270 K.

data set with over 3×10^7 entries. The total error of the neural network and q_3, $\bar{q}_3$ and $\tilde{q}_3$ is estimated and plotted in Figure 8.2. The original q_3 method performs poorly with an error of around 12%. The modification by Lechner [101], $\bar{q}_3$ is a significant improvement with an error of around 2%. Our modification, $\tilde{q}_3$ is a further small improvement to 1.51%, whilst our neural network has an error of only 0.13%.

In order to demonstrate the four classification methods in practice, we run the algorithms for a snapshot of our prism slab following TIP4P/Ice MD simulations at 270 K. A QLL is formed on the surfaces and the predictions of the four approaches is shown in Figure 8.3. Molecules that are classified as water are shown in blue and molecules that are classified as ice in red. The result illustrates the classification

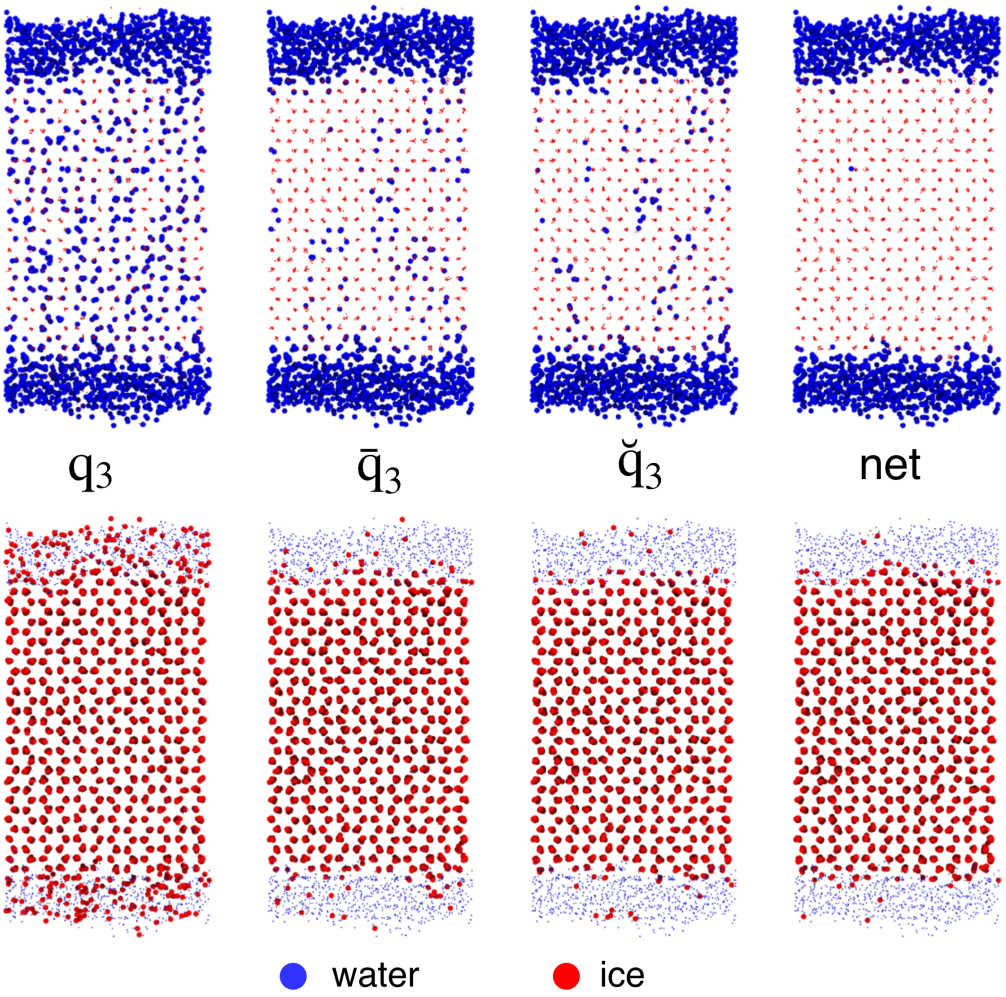


Figure 8.3: Classification of ice and water molecules by the Steinhardt and neural network algorithms for a snapshot of the prism TIP4P/Ice surface at 270 K. Molecules classified as ice are coloured red and molecules classified as water are coloured blue.

error plotted in Figure 8.2. q_3 performs poorly and misclassified a large number of bulk ice molecules as water. $\bar{q}_3$ and $\tilde{q}_3$ are a significant improvement but nonetheless misclassify a number of obvious bulk ice molecules as water. The neural network, however, misclassifies only one bulk ice molecule as water and makes only five clear mistakes in the QLL. The neural network is trained with an added surface in the y-direction, however, the snapshot in Figure 8.3 has a surface in the x-direction. The success of the network in classifying the surface QLL molecules along a new axis demonstrates that the neural network is robust and has a general understanding of surfaces. The relative coordinates used for the training ensures that the neural network is invariant to the absolute values of the system coordinates. The successful detection of a surface in a new direction suggests that the network is robust to transformations. However, rotational invariance can be ensured by applying random rotations to training input result in a more robust predictor [142].

8.3 Conclusions

In conclusion, we have developed a deep neural network which can very accurately identify ice and water phases using the coordinates of the nearest neighbours. Our approach does not rely on any knowledge of the molecular structures and is significantly more precise than the Steinhardt approach. We demonstrate that a neural network can be used to recognise different structures without the need for complex and computationally expensive symmetry functions. We offer a much simpler solution compared to the network of Geiger [139] without compromising the accuracy. Whilst the neural network developed by Geiger [139] relies on and is therefore limited by the choice of symmetry functions, ours takes in an essentially raw input data form with only one parameter, the number of nearest neighbours. Using the relative coordinates of the nearest neighbours allows our network to be applied to systems of any sizes without having to be retrained. Our network has 30 input nodes and 2 hidden layers each composed of 256 neurons. A large number of weights and biases are tuned during the training procedure as the network learns features that can be used to

recognise different structures.

The training data is produced from simple MD simulations of the pure ice and water phases. Short 1 ns NVT simulations of ice and water are performed and from only 1001 trajectory frames of each phase, an error of just 0.13% is achieved by our network on an independent test set at different temperatures. The input of our neural network takes into account only part of the second neighbour shell. Further investigation is required to characterise the dependence of the error on the number of nearest neighbours. We anticipate a decrease in accuracy as the number of nearest neighbours decreases. Once we no longer take into account the second neighbour shell, we expect a large increase in the error. A significant increase in error is also expected when going from 4 nearest neighbours to 3. The accuracy can be improved by including hydrogen atoms and the third neighbour shell.

To demonstrate our neural network in practice we run the algorithm on a snapshot of the prism slab at 270 K with a QLL surface. The network identifies bulk ice molecules with a near perfect score. When exposed to a surface along one axis during training, we demonstrate that the network develops a robust understanding of a surface which it can apply to detect surface molecules along different axes. Our deep neural network, with two hidden layers, is capable of learning complex features which characterise water and ice. The complicated task of feature engineering and extracting functions of the atomic coordinates which can be used to identify the local environment of atoms is taken care of by the neural network. No convoluted projections onto spherical harmonics or identification of symmetry functions are required. The multiple hidden layers enable the network to learn features at different abstractions. In this work, our network is given the function to identify ordered and disordered phases. Our approach can be extended to classifying ordered phases, however, further investigation is required to establish the success of the network in this task.

Chapter 9

Conclusions

The QLL of ice has been rigorously studied across a range of temperatures in this thesis using two water models. Molecular dynamics simulations indicate that the QLL size increases with temperature but that the mW QLL is over-structured. Ice and QLL molecules are distinguished using the 3^{rd}-order Steinhardt parameter but a novel neural network is identified as a future solution for classifying ice and QLL molecules. The TIP4P/Ice QLL size exhibits a logarithmic temperature dependence consistent with previous experimental and simulation works [1, 33–35]. We demonstrate that the TIP4P/Ice QLL result is predicted from a surface energy view point and using the Lindemann theory based on atomic vibrations, further validating the TIP4P/Ice description of the QLL.

The energetics of the QLL is probed in detail in this project using metadynamics. Our metadynamics simulations of QLL crystallisation are the first of their kind and provide illuminating results. The results indicate that a QLL forms with an equilibrium thickness. Melting the QLL is an activated process and there are no metastable states. The chemical potential difference of a QLL and ice molecule is computed from the simulations for TIP4P/Ice at 260 K and mW over a range of temperatures. The mW $\Delta\mu$ temperature dependence correlates with experimental results and demonstrates that whilst mW is poor at describing the equilibrium QLL structure and dynamics, it is good at describing its energetics. The metadynamics QLL crystallisation free energies show that crystallising the QLL is a spontaneous

process if the QLL is larger than its equilibrium size. Motivated by this finding, we perform unbiased MD simulations of QLL crystallisation by first artificially melting a surface.

MD simulations of an artificially melted surface reveal that QLL crystallisation occurs at a constant rate with no metastable states. The results are consistent with our metadynamics free energy profiles. The rate of QLL crystallisation is estimated from the simulations and indicate a complex temperature dependence involving an interplay of two competing effects. Building upon previous theoretical work, we describe the TIP4P/Ice rate of QLL crystallisation using the translational diffusivity, chemical potential and QLL size. The translational diffusivity and QLL size are obtained from the TIP4P/Ice MD simulations whereas the $\Delta\mu$ temperature dependence is interpolated from the mW metadynamics simulations. By modifying the Wilson-Frenkel growth equation, we predict both the functional form and numerical estimates of the TIP4P/Ice QLL crystallisation rate using only one fitting parameter: a factor which uniformly scales the mW chemical potential and takes into account that mW overestimates $\Delta\mu$ relative to TIP4P/Ice. This is a landmark result, representing the first time that the rate of QLL crystallisation has been modelled using entirely physical variables computed from independent computer simulations. Previous attempts fit the trend to arbitrary functions parameterised to the data [39, 40].

Inspired by our QLL crystallisation results, we speculate on the origin of the unexplained difference in saturation vapour pressure of supercooled water and ice. The QLL crystallisation simulations suggest a mechanism involving the diffusion of molecules into sites where they may crystallise. We speculate that a similar process is involved in evaporation and outline a hypothesis comprising an interplay of the chemical potential difference between supercooled water and ice, and the QLL structure and diffusivity. This is a novel explanation which remains speculative but merits further investigation by the community due to the important role that vapour pressure gradients play in ice growth processes within clouds.

The QLL crystallisation depends on the atomic structure and dynamics. The outer QLL structure is probed by analysing dangling bonds in the TIP4P/Ice simulations.

The orientation of dangling bonds indicate that the structure at the basal and prism QLL/vapour interfaces are closely related. Between 240 K and 265 K, the surface structure is very weakly dependent on temperature. Activation energies of translational diffusion computed from our MD simulations indicate that diffusion involves breaking two hydrogen bonds. A consistent result is obtained across the two surfaces implying that the hydrogen bonding network is similar in the two QLL surfaces. When no hydrogens are present, as in the case of mW simulations, our results show that QLL crystallisation is significantly faster as molecules are not slowed down by the reorientation of hydrogen atoms. Previous studies indicate that molecules in the QLL sublayer at the vapour interface form on average 2 hydrogen bonds, whereas molecules in deeper QLL sublayers at the QLL/ice interface form on average three hydrogen bonds [29, 38]. Transitioning into ice, molecules on average form four hydrogen bonds [29, 38]. The hydrogen bond profile impacts the molecular kinetics. We investigate this further by determining the QLL/ice interface position and extensively analysing the dynamics at the interface.

In this project we identify two robust criteria for determining the position of the TIP4P/Ice QLL/ice interface using two very different approaches. The first approach takes into account the local atomic structure whereas the second approach considers the diffusion energetics. The two criteria yield a consistent result and allow us to compute the diffusion energy barriers and rates of crossing at the QLL/ice interface from MD simulations. The results exhibit clear anisotropy at the basal and prism surfaces which depends on temperature. QLL/ice diffusion is fastest at the prism interface at low and high temperatures, and fastest at the basal interface at intermediate temperatures. As temperature increases the rate of basal and prism QLL/ice diffusion and the associated energy barriers cross-over twice. This result correlates with the two cross-overs in the morphology diagram and is a significant milestone in understanding the growth shape of ice crystals. At low and high temperatures we correctly predict thin plates whilst at intermediate temperatures we predict elongated prisms, as observed in nature. The QLL/ice diffusion energy barriers and rates of crossings results suggest that the growth shape of hexagonal

ice is determined by the QLL/ice kinetics and underlying diffusion energy barriers. The implication of our findings is that the basal and secondary prism QLL/vapour interfaces are similar and play a secondary role. This observation is in accord with our analysis of the dangling bond structure at the surface of ice.

The QLL/vapour interface is probed in detail from over 1 μs of metadynamics simulations. Our results uncover the energetics of adsorption and desorption of water molecules, and the energy barriers for diffusion from the QLL into ice. Adsorption from the vapour phase is barrierless and favourable. Desorption, however, is an activated process with a barrier equivalent to breaking two hydrogen bonds. The basal and secondary prism desorption barriers differ by less thank 1%. This result provides further testament to the equivalence of the basal and secondary prism ice surfaces that our MD results suggest. The diffusion energy barriers at the QLL/ice interface recovered from metadynamics conform with our MD estimates, and comparison with experiments indicates that the barriers correspond to the activation energy for breaking a hydrogen bond [132]. The hydrogen bond count changes by one across the QLL/ice interface [29, 38] validating a mechanism of interface diffusion involving the breaking and formation of one hydrogen bond. The anisotropy of the basal and secondary prism diffusion barriers uncovered by our work indicates that the underlying crystal structure influences this mechanism.

In this project we have rigorously identified that the QLL/ice interface is responsible for the kinetics of ice growth. Whilst the vapour phase fuels crystal growth, the ice phase is intrinsically involved and rate limiting. Further investigation into the kinetic energy and orientation of molecules diffusing at the QLL/ice interface will help uncover the molecular mechanism of ice growth and enable import progress in climate science.

Bibliography

[1] V F Petrenko and R W Whitworth. *Physics of Ice*. Clarendon Press, 1999.

[2] D. Wylie, P. Menzel, H. Woolf, and K. Strabala. Four years of global cirrus cloud statistics using hirs. *Journal of Climate*, 7(12):1972–1986, 1994.

[3] L. Kuo-Nan. Influence of cirrus clouds on weather and climate processes: A global perspective. *Monthly Weather Review*, 114(6):1167–1199, 1986.

[4] G. Stephens, S. Tsay, P. Stackhouse, and P. Flatau. The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback. *Journal of the Atmospheric Sciences*, 47(14):1742–1754, 1990.

[5] M. Schnaiter, E. Järvinen, P. Vochezer, A. Abdelmonem, R. Wagner, O. Jourdan, G. Mioche, V. N. Shcherbakov, C. G. Schmitt, U. Tricoli, Z. Ulanowski, and A. J. Heymsfield. Cloud chamber experiments on the origin of ice crystal complexity in cirrus clouds. *Atmospheric Chemistry and Physics*, 16(8):5091–5110, 2016.

[6] Thorsten Bartels-Rausch. Chemistry: Ten things we need to know about ice and snow. *Nature*, 494(7435):27–9, February 2013.

[7] M Conde, C Vega, and A Patrykiejew. The thickness of a liquid layer on the free surface of ice as obtained from computer simulation. *The Journal of Chemical Physics*, 129(1):014702, July 2008.

[8] Ukichir Nakaya. *Snow Crystals: Natural and Artificial*. Harvard University Press, 1954.

[9] T. Kobayashi. On the habit of snow crystals artificially produced at low

pressures. *Journal of the Meteorological Society of Japan Series II*, 36(5):193–208, 1958.

[10] J. Hallett and B. J. Mason. The influence of temperature and supersaturation on the habit of ice crystals grown from the vapour. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 247(1251):440–453, 1958.

[11] Kenneth G Libbrecht. The physics of snow crystals. *Reports on Progress in Physics*, 68(4):855, 2005.

[12] Kenneth G. Libbrecht. Explaining the formation of thin ice crystal plates with structure-dependent attachment kinetics. *Journal of Crystal Growth*, 258(1–2):168 – 175, 2003.

[13] Kenneth G. Libbrecht. The formation of snow crystals: subtle molecular processes govern the growth of a remarkable variety of elaborate ice structures. *American Scientist*, 95(1):52–59, 2007.

[14] V. Keller. *Ice crystal growth in a dynamic thermal diffusion chamber*. PhD thesis, Nevada Univ., Reno, November 1977.

[15] Thorsten Bartels-Rausch, Vance Bergeron, Julyan H. E. Cartwright, Rafael Escribano, John L. Finney, Hinrich Grothe, Pedro J. Gutiérrez, Jari Haapala, Werner F. Kuhs, Jan B. C. Pettersson, Stephen D. Price, Debbie J. Stokes, Giovanni Strazzulla, Erik S. Thomson, Hauke Trinks, and Nevin Uras-Aytemiz. Ice structures, patterns, and processes: A view across the icefields. *Reviews of Modern Physics*, 84:885–944, May 2012.

[16] J. D. Bernal and R. H. Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8):515–548, 1933.

[17] Steve Lower. States of matter: States of matter: Water and hydrogen bonding, February 2017.

[18] L. E. Bove, S. Klotz, A. Paciaroni, and F. Sacchetti. Anomalous proton dynamics in ice at low temperatures. *Physical Review Letters*, 103:165901, Oct 2009.

[19] Martin Chaplin. Water Structure and Science. `http://www1.lsbu.ac.uk/water/ice1h.html`. Accessed: 27/08/2014.

[20] Kenneth Libbrecht. *The Art of the Snowflake: A Photographic Album*. Voyageur Press, reissue edition edition, September 2014.

[21] D. Lamb. *Growth Rates and Habits of Ice Crystals Grown from the Vapor Phase*. PhD thesis, University of Washington, Seattle, 1970.

[22] B. J. Mason. *The physics of clouds*. Oxford University Press, 1971.

[23] James E. McDonald. Use of the electrostatic analogy in studies of ice crystal growth. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 14(5):610–620, Sep 1963.

[24] H. R. Byers. *Elements of Cloud Physics*. The University of Chicago Press, 1965.

[25] D. Lamb and Peter V. Hobbs. Growth rates and habits of ice crystals grown from the vapor phase. *Journal of the Atmospheric Sciences*, 28(8):1506–1509, 1971.

[26] Gen Sazaki, Harutoshi Asakawa, Ken Nagashima, Shunichi Nakatsubo, and Yoshinori Furukawa. How do quasi-liquid layers emerge from ice crystal surfaces? *Crystal Growth & Design*, 13(4):1761–1766, April 2013.

[27] Yimin Li and G.A. Somorjai. Surface Premelting of Ice. *Journal of Physical Chemistry C*, 111(27):9631–9637, July 2007.

[28] Gen Sazaki, Salvador Zepeda, Shunichi Nakatsubo, Makoto Yokomine, and Yoshinori Furukawa. Quasi-liquid layers on ice crystal surfaces are made up of two different phases. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4):1052–5, January 2012.

[29] Steven Neshyba, Erin Nugent, Martina Roeselová, and Pavel Jungwirth. Molecular Dynamics Study of Ice/Vapor Interactions via the Quasi-Liquid Layer. *The Journal of Physical Chemistry C*, 113(11):4597–4604, March 2009.

[30] D. Beaglehole and D. Nason. Transition layer on the surface on ice. *Surface Science*, 96(1–3):357 – 363, 1980.

[31] Yoshinori Furukawa, Masaki Yamamoto, and Toshio Kuroda. Ellipsometric study of the transition layer on the surface of an ice crystal. *Journal of Crystal Growth*, 82(4):665 – 677, 1987.

[32] H Dosch, A Lied, and J H Bilgram. Glancing-angle X-ray scattering studies of the premelting of ice surfaces. *Surface Science*, 327(1–2):145–164, 1995.

[33] I Golecki and C Jaccard. Intrinsic surface disorder in ice near the melting point. *Journal of Physics C: Solid state physics*, 4229, 1978.

[34] Hendrik Bluhm, D Frank Ogletree, Charles S Fadley, Zahid Hussain, and Miquel Salmeron. The premelting of ice studied with photoelectron spectroscopy. *Journal of Physics: Condensed Matter*, 14(8):L227, 2002.

[35] Astrid Döppenschmidt and Hans-Jürgen Butt. Measuring the thickness of the liquid-like layer on ice surfaces with atomic force microscopy. *Langmuir*, 16(16):6709–6714, 2000.

[36] C. L. Bishop, D. Pan, L. M. Liu, G. A. Tribello, A. Michaelides, E. G. Wang, and B. Slater. On thin ice: surface order and disorder during pre-melting. *Faraday Discussions*, 141:277–292, 2009.

[37] S. Picaud. Dynamics of TIP5P and TIP4P/ice potentials. *The Journal of Chemical Physics*, 125(17):174712, 2006.

[38] William Pfalzgraff, Steven Neshyba, and Martina Roeselova. Comparative molecular dynamics study of vapor-exposed basal, prismatic, and pyramidal surfaces of ice. *The journal of physical chemistry. A*, 115(23):6184–93, June 2011.

[39] J. R. Espinosa, C. Navarro, E. Sanz, C. Valeriani, and C. Vega. On the time required to freeze water. *The Journal of Chemical Physics*, 145(21):211922, 2016.

[40] Dmitri Rozmanov and Peter G. Kusalik. Anisotropy in the crystal growth of hexagonal ice, ih. *The Journal of Chemical Physics*, 137(9):094702, 2012.

[41] J Hallett. The temperature dependence of the viscosity of supercooled water. *Proceedings of the Physical Society*, 82(6):1046, 1963.

[42] Amine Dehaoui, Bruno Issenmann, and Frédéric Caupin. Viscosity of deeply supercooled water and its coupling to molecular diffusion. *Proceedings of the National Academy of Sciences*, 112(39):12020–12025, 2015.

[43] William S. Price, Hiroyuki Ide, and Yoji Arata. Self-Diffusion of Supercooled Water to 238 K Using PGSE NMR Diffusion Measurements. *The Journal of Physical Chemistry A*, 103(4):448–450, 1999.

[44] Y. Mizuno and N. Hanafusa. Studies of surface properties of ice using nuclear magnetic resonance. *Journal de Physique Colloques*, 48(C1):C1–511–C1–517, 1987.

[45] Kim Bolton and Jan B. C. Pettersson. A molecular dynamics study of the long-time ice ih surface dynamics. *The Journal of Physical Chemistry B*, 104(7):1590–1595, 2000.

[46] A. Faraone, L. Liu, C.-Y. Mou, C.-W. Yen, and S.-H. Chen. Fragile-to-strong liquid transition in deeply supercooled confined water. *The Journal of Chemical Physics*, 121(22):10843–10846, 2004.

[47] Ivan Gladich, Amrei Oswald, Natalie Bowens, Sam Naatz, Penny Rowe, Martina Roeselova, and Steven Neshyba. Mechanism of anisotropic surface self-diffusivity at the prismatic ice-vapor interface. *Physical Chemistry Chemical Physics*, 17:22947–22958, 2015.

[48] Ivan Gladich, William Pfalzgraff, Ondrej Marsalek, Pavel Jungwirth, Martina Roeselova, and Steven Neshyba. Arrhenius analysis of anisotropic surface self-diffusion on the prismatic facet of ice. *Physical Chemistry Chemical Physics*, 13:19960–19969, 2011.

[49] M. Bienfait. Roughening and surface melting transitions: consequences on crystal growth. *Surface Science*, 272(1):1 – 9, 1992.

[50] Shuji Kawada. Dielectric Dispersion and Phase Transition of KOH Doped Ice. *Journal of the Physical Society of Japan*, 32(5):1442–1442, 1972.

[51] Hiroshi Suga. A facet of recent ice sciences. *Thermochimica Acta*, 300(1):117 – 126, 1997.

[52] Hiroshi Fukazawa, Shinji Mae, Susumu Ikeda, and Okitsugu Watanabe. Proton ordering in antarctic ice observed by raman and neutron scattering. *Chemical Physics Letters*, 294(6):554 – 558, 1998.

[53] Y Wang, J.C Li, A.I Kolesnikov, S Parker, and S.J Johnsen. Inelastic neutron scattering investigation of greenland ices. *Physica B: Condensed Matter*, 276–278:282 – 283, 2000.

[54] A. D. Fortes, I. G. Wood, D. Grigoriev, M. Alfredsson, S. Kipfstuhl, K. S. Knight, and R. I. Smith. No evidence for large-scale proton ordering in antarctic ice from powder neutron diffraction. *The Journal of Chemical Physics*, 120(24):11376–11379, 2004.

[55] Chris Knight, Sherwin J. Singer, Jer-Lai Kuo, Tomas K. Hirsch, Lars Ojamäe, and Michael L. Klein. Hydrogen bond topology and the ice vii/viii and ih/xi proton ordering phase transitions. *Physical Review E*, 73:056113, May 2006.

[56] Sherwin J. Singer, Jer-Lai Kuo, Tomas K. Hirsch, Chris Knight, Lars Ojamäe, and Michael L. Klein. Hydrogen-bond topology and the ice VII/VIII and ice I*h*/XI proton-ordering phase transitions. *Physical Review Letters*, 94:135701, Apr 2005.

[57] A. Glebov, A. P. Graham, A. Menzel, J. P. Toennies, and P. Senet. A helium atom scattering study of the structure and phonon dynamics of the ice surface. *The Journal of Chemical Physics*, 112(24):11011–11022, 2000.

[58] Ding Pan, Li-Min Liu, Gareth A. Tribello, Ben Slater, Angelos Michaelides, and Enge Wang. Surface energy and surface proton order of ice I*h*. *Physical Review Letters*, 101:155703, Oct 2008.

[59] Mary Jane Shultz, Patrick J. Bisson, and Alexandra Brumberg. Best face forward: Crystal-face competition at the ice–water interface. *The Journal of Physical Chemistry B*, 118(28):7972–7980, 2014. PMID: 24784996.

[60] Ding Pan, Li-Min Liu, Gareth A Tribello, Ben Slater, Angelos Michaelides, and Enge Wang. Surface energy and surface proton order of the ice ih basal and prism surfaces. *Journal of Physics: Condensed Matter*, 22(7):074209, 2010.

[61] Lasse Makkonen. Surface melting of ice. *The Journal of Physical Chemistry B*, 101(32):6196–6200, 1997.

[62] J.M Douillard and M Henry. Calculation of surface enthalpy of solids from an ab initio electronegativity based model: case of ice. *Journal of Colloid and Interface Science*, 263(2):554 – 561, 2003.

[63] Ph. Parent, C. Laffon, C. Mangeney, F. Bournel, and M. Tronc. Structure of the water ice surface studied by x-ray absorption spectroscopy at the O K-edge. *The Journal of Chemical Physics*, 117(23):10842–10851, 2002.

[64] Rossend Rey, Klaus B. Møller, and James T. Hynes. Hydrogen bond dynamics in water and ultrafast infrared spectroscopy. *The Journal of Physical Chemistry A*, 106(50):11993–11996, 2002.

[65] C. P. Lawrence and J. L. Skinner. Vibrational spectroscopy of HOD in liquid D2O. III. Spectral diffusion, and hydrogen-bonding and rotational dynamics. *The Journal of Chemical Physics*, 118(1):264–272, 2003.

[66] M. Alejandra Sánchez, Tanja Kling, Tatsuya Ishiyama, Marc-Jan van Zadel, Patrick J. Bisson, Markus Mezger, Mara N. Jochum, Jenée D. Cyran, Wilbert J. Smit, Huib J. Bakker, Mary Jane Shultz, Akihiro Morita, Davide Donadio, Yuki Nagata, Mischa Bonn, and Ellen H. G. Backus. Experimental and theoretical evidence for bilayer-by-bilayer surface melting of crystalline ice. *Proceedings of the National Academy of Sciences*, 114(2):227–232, 2017.

[67] Geert-Jan Kroes. Surface melting of the (0001) face of TIP4P ice. *Surface Science*, 275(3):365 – 382, 1992.

[68] M. Watkins, D. Pan, E. G. Wang, A. Michaelides, J. VandeVondele, and B. Slater. Large variation of vacancy formation energies in the surface of crystalline ice. *Nature Materials*, 10(10):794–798, 10 2011.

[69] Xing Wei, Paulo B. Miranda, and Y. R. Shen. Surface vibrational spectroscopic study of surface melting of ice. *Physical Review Letters*, 86:1554–1557, Feb 2001.

[70] Harutoshi Asakawa, Gen Sazaki, Ken Nagashima, Shunichi Nakatsubo, and Yoshinori Furukawa. Two types of quasi-liquid layers on ice crystals are formed kinetically. *Proceedings of the National Academy of Sciences*, 113(7):1749–1753, 2016.

[71] Ken-ichiro Murata, Harutoshi Asakawa, Ken Nagashima, Yoshinori Furukawa, and Gen Sazaki. Thermodynamic origin of surface melting on ice crystals. *Proceedings of the National Academy of Sciences*, 113(44):E6741–E6748, 2016.

[72] David T. Limmer. Closer look at the surface of ice. *Proceedings of the National Academy of Sciences*, 113(44):12347–12349, 2016.

[73] Matteo Salvalaglio, Thomas Vetter, Federico Giberti, Marco Mazzotti, and Michele Parrinello. Uncovering molecular details of urea crystal growth in the presence of additives. *Journal of the American Chemical Society*, 134(41):17221–33, October 2012.

[74] Adri C. T. van Duin, Siddharth Dasgupta, Francois Lorant, and William A. Goddard. Reaxff:â a reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41):9396–9409, 2001.

[75] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[76] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993.

[77] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327 – 341, 1977.

[78] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[79] Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, 1980.

[80] Daan Frenkel and Berend Smit, editors. *Understanding Molecular Simulation: From Algorithms to Applications.* Academic Press, Inc., Orlando, FL, USA, 1st edition, 1996.

[81] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[82] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31:1695–1697, Mar 1985.

[83] Philippe H. Hünenberger. Thermostat algorithms for molecular dynamics simulations. *Advances in Polymer Science*, 173:105–149, 2005.

[84] Alessandro Laio, Antonio Rodriguez-Fortea, Francesco Luigi Gervasio, Matteo Ceccarelli, and Michele Parrinello. Assessing the accuracy of metadynamics. *The Journal of Physical Chemistry B*, 109(14):6714–6721, 2005. PMID: 16851755.

[85] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002.

[86] Francesco Luigi Gervasio, Alessandro Laio, and Michele Parrinello. Flexible docking in solution using metadynamics. *Journal of the American Chemical Society*, 127(8):2600–2607, 2005. PMID: 15725015.

[87] Federico Comitani. *Computational Modelling of Activation Mechanisms in Ligand Gated Ion Channels*. PhD thesis, King's College London, 2016.

[88] J L F Abascal and C Vega. A general purpose model for the condensed phases of water: TIP4P/2005. *The Journal of Chemical Physics*, 123(23):234505, December 2005.

[89] Carlos Vega, José L F Abascal, M M Conde, and J L Aragones. What ice can teach us about water interactions: a critical comparison of the performance of different water models. *Faraday Discussions*, 141:251, 2009.

[90] J L F Abascal, E Sanz, R García Fernández, and C Vega. A potential model for the study of ices and amorphous water: TIP4P/Ice. *The Journal of Chemical Physics*, 122(23):234511, June 2005.

[91] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91(24):6269–6271, 1987.

[92] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

[93] Pekka Mark and Lennart Nilsson. Structure and dynamics of the tip3p, spc,

and spc/e water models at 298 k. *The Journal of Physical Chemistry A*, 105(43):9954–9960, 2001.

[94] Valeria Molinero and Emily B Moore. Water modeled as an intermediate element between carbon and silicon. *The Journal of Physical Chemistry B*, 113(13):4008–16, April 2009.

[95] Aleks Reinhardt and Jonathan P K Doye. Free energy landscapes for homogeneous nucleation of ice for a monatomic water model. *The Journal of chemical physics*, 136(5):054501, February 2012.

[96] V. Buch, P. Sandler, and J. Sadlej. Simulations of H2O Solid, Liquid, and Clusters, with an Emphasis on Ferroelectric Ordering Transition in Hexagonal Ice. *The Journal of Physical Chemistry B*, 102(44):8641–8653, 1998.

[97] C. Vega and E. de Miguel. Surface tension of the most popular models of water by using the test-area simulation method. *The Journal of Chemical Physics*, 126(15), 2007.

[98] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1 – 19, 1995.

[99] Pieter Rein ten Wolde, Maria J. Ruiz-Montero, and Daan Frenkel. Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling. *The Journal of Chemical Physics*, 104(24):9932, 1996.

[100] Emily B Moore, Ezequiel de la Llave, Kai Welke, Damian a Scherlis, and Valeria Molinero. Freezing, melting and structure of ice in a hydrophilic nanopore. *Physical chemistry chemical physics : PCCP*, 12(16):4124–34, April 2010.

[101] Wolfgang Lechner and Christoph Dellago. Accurate determination of crystal structures based on averaged local bond order parameters. *The Journal of chemical physics*, 129(11):114707, September 2008.

[102] Aleks Reinhardt, Jonathan P K Doye, Eva G Noya, and Carlos Vega. Local order parameters for use in driving homogeneous ice nucleation with all-atom

models of water. *The Journal of chemical physics*, 137(19):194504, November 2012.

[103] Wikimedia Commons. Spherical Harmonics. `http://en.wikipedia.org/wiki/Spherical_harmonics`.

[104] Swetlana Jungblut and Christoph Dellago. Crystallization of a binary Lennard-Jones mixture. *The Journal of chemical physics*, 134(10):104501, March 2011.

[105] Paul Steinhardt, David Nelson, and Marco Ronchetti. Bond-orientational order in liquids and glasses. *Physical Review B*, 28(2):784–805, July 1983.

[106] E Sanz, C Vega, J R Espinosa, R Caballero-Bernal, J L F Abascal, and C Valeriani. Homogeneous ice nucleation at moderate supercooling from molecular simulation. *Journal of the American Chemical Society*, 135(40):15008–17, October 2013.

[107] H. Nada and Y. Furukawa. Anisotropy in structural phase transitions at ice surfaces: a molecular dynamics study. *Applied Surface Science*, 121-122:445–447, November 1997.

[108] Charusita Chakravarty, Pablo G Debenedetti, and Frank H Stillinger. Lindemann measures for the solid-liquid phase transition. *The Journal of chemical physics*, 126(20):204508, May 2007.

[109] Kaiwang Zhang, G Malcolm Stocks, and Jianxin Zhong. Melting and premelting of carbon nanotubes. *Nanotechnology*, 18(28):285703, July 2007.

[110] V. Buch and J.P. Devlin. *Water in Confining Geometries*. Springer Series in Cluster Physics. Springer Berlin Heidelberg, 2013.

[111] Alenka Luzar and David Chandler. Hydrogen-bond kinetics in liquid water. *Nature*, 379(6560):55–57, 01 1996.

[112] Jill Tomlinson-Phillips, Joel Davis, Dor Ben-Amotz, Daniel Spångberg, Ljupčo Pejov, and Kersti Hermansson. Structure and Dynamics of Water Dangling

OH Bonds in Hydrophobic Hydration Shells. Comparison of Simulation and Experiment. *The Journal of Physical Chemistry A*, 115(23):6177–6183, 2011. PMID: 21413757.

[113] Gareth A. Tribello, Massimiliano Bonomi, Davide Branduardi, Carlo Camilloni, and Giovanni Bussi. PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, 185(2):604 – 613, 2014.

[114] Flaviu S. Cipcigan, Vlad P. Sokhan, Andrew P. Jones, Jason Crain, and Glenn J. Martyna. Hydrogen bonding and molecular orientation at the liquid-vapour interface of water. *Phys. Chem. Chem. Phys.*, 17:8660–8669, 2015.

[115] Thomas D. Kuhne, Tod A. Pascal, Efthimios Kaxiras, and Yousung Jung. New insights into the structure of the vapor/water interface from large-scale first-principles simulations. *The Journal of Physical Chemistry Letters*, 2(2):105–113, 2011. PMID: 26295528.

[116] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press Inc., 1987.

[117] S. J. Suresh and V. M. Naik. Hydrogen bond thermodynamic properties of water from dielectric constant data. *The Journal of Chemical Physics*, 113(21):9727–9732, 2000.

[118] M. A. Carignano, P. B. Shepson, and I. Szleifer *. Molecular dynamics simulations of ice growth from supercooled water. *Molecular Physics*, 103(21-23):2957–2967, 2005.

[119] O.B. Nasello, S. Navarro de Juarez, and C.L. Di Prinzio. Measurement of self-diffusion on ice surface. *Scripta Materialia*, 56(12):1071 – 1073, 2007.

[120] Dmitri Rozmanov and Peter G. Kusalik. Transport coefficients of the tip4p-2005 water model. *The Journal of Chemical Physics*, 136(4):044507, 2012.

[121] Y. Saito. *Statistical Physics of Crystal Growth*. World Scientific Books, 1996.

[122] Kenneth G. Libbrecht. Toward a comprehensive model of snow crystal growth: 3. the correspondence between ice growth from water vapor and ice growth from liquid water. *arXiv*, July 2014.

[123] Amir Haji-Akbari and Pablo G. Debenedetti. Direct calculation of ice homogeneous nucleation rate for a molecular model of water. *Proceedings of the National Academy of Sciences*, 112(34):10582–10588, 2015.

[124] T. Kuroda and R. Lacmann. Growth kinetics of ice from the vapour phase and its growth forms. *Journal of Crystal Growth*, 56:189–205, 1982.

[125] J. S. Wettlaufer, M. G. Worster, L. A. Wilen, and J. G. Dash. A theory of premelting dynamics for all power law forces. *Physical Review Letters*, 76:3602–3605, May 1996.

[126] H. R. Pruppacher. Interpretation of experimentally determined growth rates of ice crystals in supercooled water. *The Journal of Chemical Physics*, 47(5):1807–1813, 1967.

[127] David Turnbull. Under what conditions can a glass be formed? *Contemporary Physics*, 10(5):473–488, 1969.

[128] University Corporation of Atmospheric Research: COMET Program. Winter precipitation processes: The ice-crystal process, March 2017.

[129] Kenneth Libbrecht. Growth rates of the principal facets of ice between $-10°$C and $-40°$C. *Journal of Crystal Growth*, 247(3–4):530 – 540, 2003.

[130] Steven Neshyba, Jonathan Adams, Kelsey Reed, Penny M. Rowe, and Ivan Gladich. A quasi-liquid mediated continuum model of faceted ice dynamics. *Journal of Geophysical Research: Atmospheres*, 121(23):14,035–14,055, 2016. 2016JD025458.

[131] Xiangrui Kong, Panos Papagiannakopoulos, Erik S. Thomson, Nikola Marković, and Jan B. C. Pettersson. Water accommodation and desorption kinetics on

ice. *The Journal of Physical Chemistry A*, 118(22):3973–3979, 2014. PMID: 24814567.

[132] Jared D. Smith, Christopher D. Cappa, Kevin R. Wilson, Benjamin M. Messer, Ronald C. Cohen, and Richard J. Saykally. Energetics of hydrogen bond network rearrangements in liquid water. *Science*, 306(5697):851–853, 2004.

[133] Rainer Feistel and Wolfgang Wagner. Sublimation pressure and sublimation enthalpy of H2O ice Ih between 0 and 273.16 K. *Geochimica et Cosmochimica Acta*, 71(1):36 – 45, 2007.

[134] John Vieceli, Martina Roeselová, and Douglas J Tobias. Accommodation coefficients for water vapor at the air/water interface. *Chemical Physics Letters*, 393(1–3):249 – 255, 2004.

[135] Enrique R. Batista, Patrick Ayotte, Ante Bilić, Bruce D. Kay, and Hannes Jónsson. What determines the sticking probability of water molecules on ice? *Physical Review Letters*, 95:223201, Nov 2005.

[136] D. R. Haynes, N. J. Tro, and S. M. George. Condensation and evaporation of water on ice surfaces. *The Journal of Physical Chemistry*, 96(21):8502–8509, 1992.

[137] T. W. Choularton and J. Latham. Measurements of the deposition coefficient for ice, and its application to cirrus seeding. *Quarterly Journal of the Royal Meteorological Society*, 103(436):307–318, 1977.

[138] J. Skrotzki, P. Connolly, M. Schnaiter, H. Saathoff, O. Möhler, R. Wagner, M. Niemand, V. Ebert, and T. Leisner. The accommodation coefficient of water molecules on ice: cirrus cloud studies at the aida simulation chamber. *Atmospheric Chemistry and Physics*, 13(8):4451–4466, 2013.

[139] Philipp Geiger and Christoph Dellago. Neural networks for local structure detection in polymorphic systems. *The Journal of Chemical Physics*, 139(16):164105, 2013.

[140] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 10 1986.

[141] Jimmy Lei Ba Diederik P. Kingma. Adam: A method for stochastic optimization. *3rd International Conference for Learning Representations*, May 2015.

[142] P. Y. Simard, D. Steinkraus, and J. C. Platt. Best practices for convolutional neural networks applied to visual document analysis. In *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.*, pages 958–963, Aug 2003.